Genetic Differences in Myelodysplastic Syndrome and Clonal Cytopenia of Undetermined Significance

KANG, YEHYUN

Department of Medical Science

Graduate School

Yonsei University

Genetic Differences in Myelodysplastic Syndrome and Clonal Cytopenia of Undetermined Significance

Advisor LEE, SEUNG-TAE

A Master's Thesis Submitted
to the Department of Medical Science
and the Committee on Graduate School
of Yonsei University in Partial Fulfillment of the
Requirements for the Degree of
Master of Medical Science

KANG, YEHYUN

June 2025

Genetic Differences in Myelodysplastic Syndrome and Clonal Cytopenia of Undetermined Significance

This Certifies that the Master's Thesis of KANG, YEHYUN is Approved

Committee Chair	LEE, KYUNG-A
Committee Member	LEE, SEUNG-TAE
Committee Member	KIM, JI EUN

Department of Medical Science
Graduate School
Yonsei University
June 2025

ACKNOWLEDGEMENTS

Every achievement has been supported by those who believe in me. I am truly grateful for their encouragement, guidance, and kindness.

First, I sincerely appreciate my advisor, Professor Seung-Tae Lee, for his continuous support and guidance throughout my graduate studies. His insightful advice and expertise have been instrumental in shaping this research and helping me grow as a researcher.

I am also deeply grateful to my thesis committee members, Professor Kyung-A Lee and Professor Ji Eun Kim, for their valuable time, thoughtful feedback, and suggestions, which have greatly enhanced the quality of this work.

My sincere thanks go to Professor Jong Rak Choi, for his continued interest in my work, which has been a constant source of motivation. I also greatly appreciate the support of Professor Saeam Shin, for her thoughtful attention and feedback, which helped me refine my ideas, learn, and grow throughout this process.

I would also like to thank Yu Jeong Choi for the helpful contributions and valuable ideas related to this research. Additionally, I am thankful to Professor Dongju Won, and Yu Jin Park for their advice and insightful suggestions on my analysis.

Special thanks to my lab colleagues—Hyeonah Lee, Miri Park, Taek Gyu Lee, Eunju Yeom, and Seoyoung Lim—for their endless support and encouragement. Their companionship made the long hours in the lab more enjoyable and having them by my side was a great source of strength throughout this journey. Thanks for all the laughs and great memories during my time in the lab.

I am also thankful to my friends, who have always supported and encouraged me in their own ways, each bringing me comfort, strength, and laughter when I needed it most.

Lastly, I am deeply grateful to my family for their love, patience and endless encouragement. Their constant belief in me made this journey possible, and I could not have achieved it without them.

I will always remember and be grateful for the help and support I've received. I promise to continue growing as a better researcher and a better person. Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Genetic Differences in Myelodysplastic Syndrome and Clonal Cytopenia of Undetermined Significance

Clonal hematopoiesis of undetermined significance (CCUS) is characterized by persistent cytopenia of unexplained origin accompanied by clonal hematopoiesis (CH). Understanding its genetic landscape and distinction from myelodysplastic syndrome (MDS) is crucial for improving diagnosis and predicting disease progression. This study investigates genetic differences and their clinical implications in CCUS and MDS.

We analyzed a total of 447 patients with persistent cytopenia who underwent next-generation sequencing (NGS) at our institution. Hybrid capture-based NGS targeted 30 or 531 hematologic cancer-related genes (Dxome, Seongnam, Korea) on the NextSeq 550Dx platform (Illumina, CA, USA)[1].

Among these, 238 patients were diagnosed with MDS and 49 with CCUS. More than two-thirds of MDS patients (74.8%, 178/238) harbored at least one tier 1/2 somatic variant. The average mutation count per person was slightly higher in MDS (1.79; range, 0-11) than in CCUS (1.57; range, 0-8), and the average of variant allele frequency (VAF) was also significantly higher in MDS (21.27% vs. 7.22%, $p < 0.001$). *ASXL1*, *SF3B1*, *RUNX1*, and *EZH2* mutations were more common in MDS, whereas *DNMT3A* mutations were predominant in CCUS. Functional categorization of 30 genes revealed distinct mutation patterns with DNA methylation-related mutations significantly enriched in CCUS and splicing factor mutations more frequent in MDS ($p < 0.001$).

These findings highlight key genetic differences between CCUS and MDS, providing insights into their distinct pathophysiology and progression risk. While our study enhances the understanding of CH and its transition to hematologic malignancies, longer follow-up is needed to assess disease progression. Future research with larger cohorts and extended follow-up is essential for refining diagnostic and prognostic strategies, improving risk stratification and patient management.

Key words : myelodysplastic syndrome, clonal cytopenia of undetermined significance, next-generation-sequencing

1. Introduction

Clonal hematopoiesis (CH) is an age-associated condition characterized by the presence of hematopoietic stem or progenitor cells that carry somatic mutations[2]. These mutations provide a selective growth advantage to the affected cells, leading to their expansion and the accumulation of clonal populations in the blood. CH can occur without clear symptoms, and in many cases, individuals with CH do not meet the diagnostic criteria for a hematologic malignancy[3]. However, these somatic mutations are often detected in peripheral blood, suggesting a tendency to develop into hematologic disorders.

Previously considered an incidental finding, CH has gained attention due to its possible involvement in both hematologic and non-hematologic diseases[4]. Background research now suggests that CH is correlated with an elevated risk of hematologic malignancies, cardiovascular diseases, and inflammatory conditions[5]. The likelihood of CH increases with age[6], with about 10% of people over 70 showing detectable clonal mutations in their blood[7]. Understanding CH is crucial, not just because it may lead to hematologic cancers, but also because it could play a role in the development of other systemic health conditions.

Recent progress in genomic testing, especially through next-generation sequencing (NGS)-based methods, has greatly improved our knowledge of CH. With the increased detection sensitivity of NGS, somatic variants have been identified in the blood of individuals who do not meet the diagnostic criteria for hematologic malignancies.[8]. This resulted in the identification of several precursor conditions, including clonal hematopoiesis of indeterminate potential (CHIP), idiopathic dysplasia of undetermined significance (IDUS), idiopathic cytopenia of undetermined significance (ICUS), and clonal cytopenia of undetermined significance (CCUS)[9]. These conditions have gained interest as they may represent early stages in the development of hematologic malignancies.

CH-associated mutations commonly involve genes regulating epigenetic modification, DNA methylation, and hematopoietic self-renewal, such as *DNMT3A*, *TET2*, and *ASXL1*[2]. These mutations give a selective benefit to hematopoietic clones, leading their expansion over time. Despite its high prevalence in aging populations, it does not always progress to clinically significant malignancies. Therefore, further research is required to elucidate the factors that contribute to disease progression and the transition from clonal hematopoiesis to hematologic malignancies.

To date, hematologic malignancies have been understood to develop through the stepwise accumulation of genetic alterations, ultimately resulting in uncontrolled proliferation and impaired cellular differentiation. More recent insights into CH have refined this perspective, revealing that clonal mutations may exist in a premalignant state for years before progressing to detectable disease. As shown in **Figure 1**, CH can advance from CHIP to CCUS, MDS, and eventually AML. However, this progression is not universal, as some cases remain stable over time without advancing to malignancy. Each stage is characterized by additional genetic and clinical abnormlities, including cytopenia, dysplasia, and cytogenetic and molecular changes[10]. Understanding thistransition is crucial for identifying high-risk individuals and potential intervention strategies.

CHIP is defined as the presence of clonal hematopoiesis without cytopenia or other hematologic abnormalities. It is often detected incidentally in healthy individuals undergoing genetic screening. In contrast, CCUS is characterized by persistent cytopenia accompanied by clonal hematopoiesis in the absence of myelodysplastic syndrome (MDS)-defining criteria[11]. Differentiating these conditions is critical for patient management, as CCUS has a significantly higher risk of progression to MDS or other myeloid malignancies compared to CHIP.

The primary distinction between CHIP and CCUS lies in their clinical presentation. While CHIP is largely asymptomatic and identified through genetic screening, CCUS often appeareswith clinically significant cytopenia, requiring hematologic evaluation. The presence of CH in both conditions suggests a shared origin but, CCUS patients exhibit a higher burden of mutations in genes associated with hematopoietic dysregulation[12]. This implies that CCUS may represent an early stage of clonal evolution leading myeloid malignancy.

Given the higher risk associated with CCUS, distinguishing it from CHIP is crucial for clinical decision-making. While CHIP may not require immediate intervention, CCUS patients benefit from close monitoring and risk stratification to identify those at greatest risk for disease progression[12]. The ability to differentiate these conditions through genetic and clinical parameters will improve patient management and guide therapeutic strategies.

The recent recognition of CCUS as a distinct entity in the 2022 International Consensus Classification (ICC) and World Health Organization (WHO) classification systems underscores its clinical importance[13]. As a condition that bridges benign clonal hematopoiesis and detectable malignancy, CCUS represents a critical stage where early intervention may prevent disease

progression.

The diagnosis of CCUS relies on a combination of clinical, hematologic, and molecular criteria. Persistent unexplained cytopenia is the main feature of CCUS, but other possible causes, such as nutritional deficiencies, autoimmune disease, and bone marrow failure syndromes, must be ruled out before making a diagnosis. Finding somatic mutations in important key hematopoietic genes helps confirm the diagnosis, and next-generation sequencing (NGS) is becoming a useful tool in clinical practice.

Management strategies for CCUS are still evolving. Considering heterogeneity of the condition, treatment approaches differ from watchful waiting to early therapeutic intervention in high-risk cases. Patients having high-risk mutations may benefit from closer monitoring and potential enrollment in clinical trials.

Studies have shown that certain genetic profiles in CCUS patients increase the risk of progression to myeloid neoplasms[14 15]. Patients with CCUS may also experience significant clinical symptoms such as fatigue or frequent infections due to anemia or leukopenia. Given its premalignant potential, distinguishing CCUS from MDS is essential for early diagnosis, prognosis, and clinical management. These issues can significantly affect the patient's quality of life and necessitate careful monitoring and management. In this study, we aimed to compare the genetic profiles of CCUS and MDS patients to identify mutations associated with clinical features and potential prognostic significance.

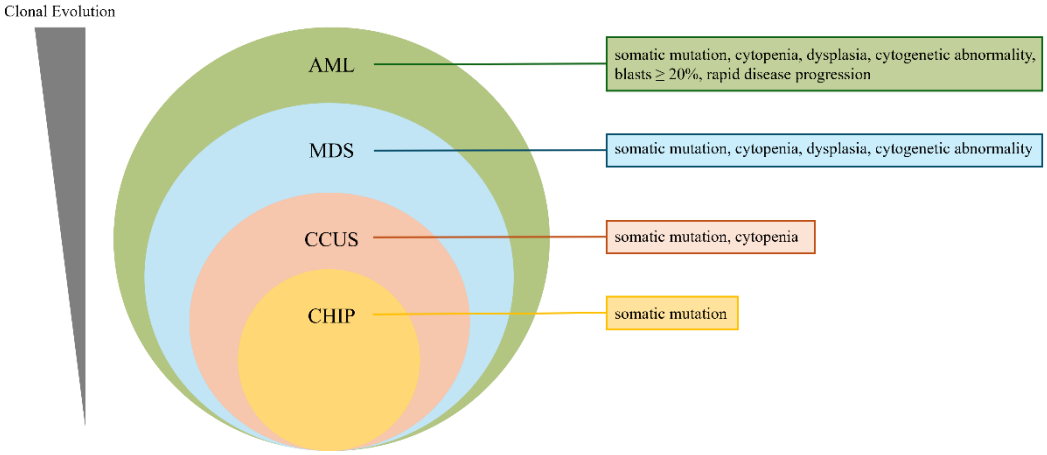

Figure 1. Stepwise progression of clonal hematopoiesis of indetermined potential (CHIP) to myeloid malignancies. Clonal hematopoiesis (CH) may lead to the development of CHIP, characterized by the presence of somatic mutations[16]. Clonal cytopenia of undetermined significance (CCUS) is characterized by the presence of somatic mutations with ongoing cytopenia[12]. Myelodysplastic syndrome (MDS) is distinguished by additional features like dysplasia and cytogenetic abnormalities. In some cases, MDS may progress to MDS-related AML which is defined by the presence of ≥20% blasts and rapid disease progression[17].

2. Materials and Methods

2.1. Patients and Samples

This retrospective descriptive study included patients with persistent cytopenia(s) who underwent next-generation sequencing (NGS) between October 2018 and June 2024 at Yonsei University Severance Hospital (Seoul, Korea). Cytopenias were defined according to the 5th edition of the World Health Organization (WHO) Classification of Hematolymphoid Tumors as follows: hemoglobin (Hb) < 13 g/dL in males and < 12 g/dL in females for anemia, absolute neutrophil count (ANC) < 1.8×10^9/L for leukopenia, and platelets < 150×10^9/L for thrombocytopenia[18]. Patients were considered transfusion dependent (TD) if they required at least 2 units of RBCs or platelets on average over 4 weeks, or 4 units or more within 8 weeks[19].

According to the same WHO classification, clonal cytopenia of undetermined significance (CCUS) is characterized by persistent cytopenias lasting for at least four months, in the absence of morphologic dysplasia in the bone marrow, but with the presence of one or more somatic mutations associated with myeloid neoplasm.

In contrast, myelodysplastic neoplasms (MDS) are diagnosed in patients with persistent cytopenias who exhibit morphologic evidence of dysplasia affecting ≥ 10% of cells in at least one hematopoietic lineage, or who harbor MDS-defining cytogenetic or molecular abnormalities such as del(5q), -7/del(7q), complex karyotypes, or *SF3B1* mutations with ring sideroblasts.

MDS was classified into 6 subtypes according to the WHO in 2022. Following that classification, we also grouped our MDS patients into 6 subtypes: MDS with biallelic *TP53* inactivation (MDS-bi*TP53*), MDS with increased blasts (MDS-IB), MDS with low blasts and isolated 5q deletion (MDS-5q), MDS with low blasts and *SF3B1* mutation (MDS-*SF3B1*), MDS, hypoplastic (MDS-h), and MDS with low blasts (MDS-LB)[18].

Bone marrow cellularity was defined as normocellular if it was within ±20% of (100 minus the patient's age), hypercellular if it exceeded this range by more than 20%, and hypocellular if it was more than 20% below this range.

Residual bone marrow samples from enrolled patients were collected during their routine clinical testing. The samples were processed to extract genomic DNA (gDNA) using the chemagic™ DNA Blood200 kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. Genomic

DNA was quantified using the Qubit Broad Range (BR) dsDNA kit (Invitrogen, Carlsbad, CA). Approximately 200ng of bone marrow-derived gDNA was processed using the Twist Library Preparation EF Kit (Twist Bioscience, San Francisco, USA). DNA fragmentation was carried out for 18 minutes at 32℃, followed by enzyme inactivation at 65℃ for 30 minutes. Hybrid capture-based NGS was performed using custom probes targeting 30 (MRD30) or 531 (Hema500) hematologic cancer-related genes (Dxome, Seongnam, Korea) and NextSeq 550Dx Instrument (Illumina, CA, USA) as shown in **Table1**.

Figure2 shows a schematic of the study workflow. A total 447 patients were included in this study, and there were 238 MDS and 49 CCUS patients included. Clinical data, including blast percentage, dysplasia status, and complete blood count (CBC) results were retrospectively collected from the elctronic medical records (EMR) system. This research received approval from the Yonsei University Severance Hospital Institutional Review Board (IRB No.4-2025-0397).

2.2. Data processing

Raw sequencing data were processed using the bcl2fastq conversion software (Illumina) to generate FASTQ files. Bioinformatics pipelines consist of demultiplexing, read alignment, deduplication, base recalibration and variant calling. Adapter sequences were removed, and reads were aligned to the hg19 reference genome using the Burrows-Wheeler Aligner (BWA; version 0.7.12; Wellcome Trust Sanger Institute, Cambridge, UK) [20]. Variant calling was performed using the PiSeq algorithm (Dxome), an error-correction pipeline that improves accuracy by enabling more precise localization of read group mappings, thereby enhancing the detection of single nucleotide variants (SNVs) and small insertions/deletions (indels) (**Figure 3**)[21,22]. Variant annotation was carried out with DxSeq software (Dxome) using public databases like gnomAD, 1000 Genomes, ClinVar, COSMIC and dbSNP.

Somatic variants were classified based on multiple guidelines, including those from the Association for Molecular Pathology (AMP), American Society of Clinical Oncology (ASCO), and College of American Pathologists (CAP)[23]. Germline variants were assessed according to the American College of Medical Genetics and Genomics (ACMG) and AMP criteria[24]. To minimize false positives, all somatic variants were re-examined with Integrative Genomics Viewer (IGV)[25]. Germline variants were excluded by cross-referencing with matched germline data obtained from

skin biopsy or buccal swab.

Table 1. Gene list of customed MRD30 panel

MRD30	
ASXL1	*MPL*
BCOR	*NPM1*
CALR	*NRAS*
CEBPA	*PHF6*
DDX41	*PTPN11*
DNMT3A	*RAD21*
ETV6	*RUNX1*
EZH2	*SETD2*
FLT3	*SF3B1*
IDH1	*SRSF2*
IDH2	*STAG2*
JAK2	*TET2*
KIT	*TP53*
KMT2D	*U2AF1*
KRAS	*WT1*

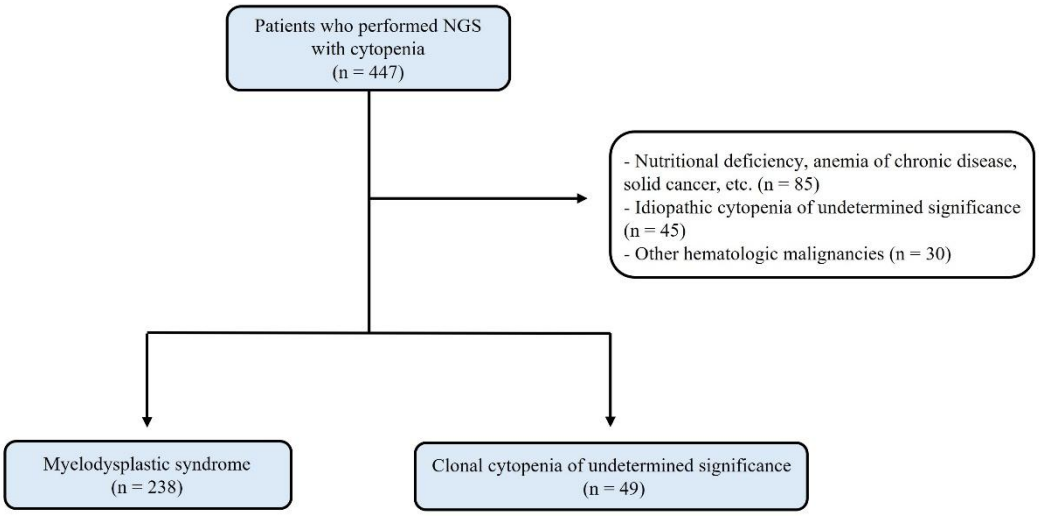

Figure 2. Flowchart of the patients enrolled in our study.

2.3. Data reanalysis

Initially, routine NGS analysis was performed using the Hema500 panel with an exclusion of variant allele frequency (VAF) below 5% variants until June 2023. However, given the concern that focusing simply on variants with a VAF of ≥2% might overlook clinically significant mutations, we sought to enhance the sensitivity of our approach by reanalyzing previous data without a strict VAF cutoff. To achieve this, we incorporated the MRD30 panel, which consists of 30 key genes associated with hematologic malignancies, for both the reanalysis and subsequent analyses. This adjustment allowed for a more comprehensive assessment of low-frequency mutations, improving the detection of clinically relevant variants.

Initial analysis was performed using a gene set of pathogenic, likely pathogenic, and tier 1 and 2 variants from the Hema500 targeted panel sequencing, followed by a reanalysis of the data using the MRD30 targeted panel. The reanalysis utilized the Piseq algorithm to enhance the accuracy of variant calling by reducing sequencing noise and improving the sensitivity for low VAF mutations. Notably, the MRD30 panel has a limit of detection (LoD) of 0.25%, enabling detection of very low-frequency variants [26]. In our study, reanalysis was performed without applying a VAF threshold, in order to capture low-level mutations that might be missed using conventional cutoffs. Given recent reports that a VAF cutoff of 5% may fail to detect fast-expanding clones and certain mutations[27], this approach allowed us to provide more sensitive and comprehensive results, establishing a key distinction from prior research.

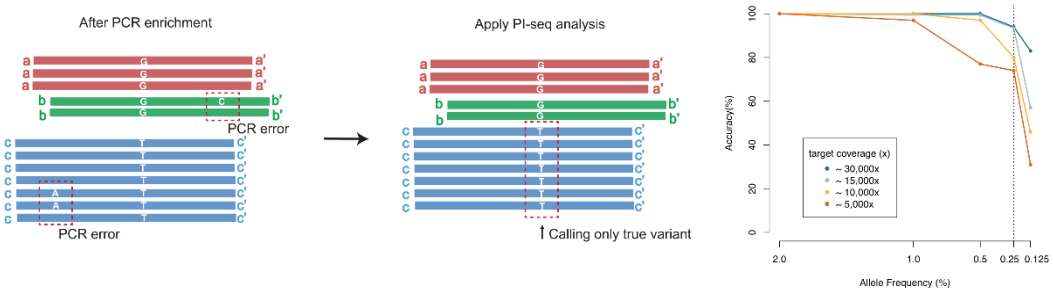

Figure 3. Positional indexing (Piseq) algorithm. This algorithm was used to correct errors occurring during the NGS (Next generation sequencing) process and to improve the accuracy of low VAF (Variant allele frequency) analysis.

2.4. Interpretation criteria

Somatic variants were classified into tiers 1,2, or 3 based on their oncogenicity and clinical significance, following the guidelines of the American Society of Clinical Oncology (ASCO), Association for Molecular Pathology (AMP), and College of American Pathologists (CAP).

Tier 1 variants included pathogenic or likely pathogenic mutations that are well-recognized for their clinical relevance. Tier 2 variants included mutations with potential clinical implications, while Tier 3 variants represented variants of uncertain significance (VUS), where there was insufficient data to find out their pathogenic potential. In this study, only somatic variants classified as Tier 1 or Tier 2 were included.

To minimize false positives, all somatic variants were manually reviewed using the Integrative Genomics Viewer (IGV)[25]. Germline variants were excluded by cross-referencing with matched germline data obtained from buccal swabs. The most common causes of false positives included low base call quality, low mapping quality, detection in multiple samples within the same batch, strand bias, and mutations occurring in repeat or homopolymer regions. To ensure accuracy, these regions were carefully reviewed, and false-positive variants were excluded, leaving only true positive variants.

In our analysis, germline variants were generally excluded except for *DDX41* gene mutations due to their established association with hematologic malignancies. DDX41 is a germline predisposition gene associated with myeloid neoplasms, particularly myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML)[28] [29]. Pathogenic germline variants in *DDX41* increase the risk of developing leukemia, and subsequent acquisition of somatic mutations in the same gene can further drive disease progression. Given these findings, *DDX41* germline variants were retained in our study to allow for a more comprehensive assessment of their role in hematologic malignancy development.

2.5. Longitudinal analysis

We investigated the survival status of CCUS and MDS patients until May 2025. Additionally, we examined disease progression from CCUS to MDS and from MDS to AML. Disease progression was defined based on the diagnostic classification at the time of follow-up, determined through bone marrow examination. Complete blood count (CBC) data were extracted from electronic medical records (EMRs) at the time of routine follow-up visits.

2.6. Meta-analysis

To enhance the reliability of our findings, we conducted a meta-analysis to integrate results from similar studies. Studies were searched by Cochrane Library, MEDLINE, PubMed and google scholar using keywords like "*MDS mutation*", "*CCUS progression*", "*CCUS and MDS*", and "*clonal hematopoiesis*". Original review articles after 2010 were included and available genomic data (e.g., variant allele frequencies (VAF), identified mutations, mutation numbers, 95% CI, p-value) were collected. We reviewed the sequencing methods used in those studies (e.g., target sequencing, whole-genome sequencing) and identified overlapping genes with our MRD30 panel. Commonly studied genes across these publications were selected for inclusion in our integrative analysis.

R-based meta-analysis packages such as "meta", "metafor" was used[30]. Random effect model was used because the studies showed no considerable heterogeneity by testing Q statistic or I^2. We checked for the publication bias with Funnel plot and for the result showing publication bias, we used "trimfill" in metafor library to adjust the publication bias. By this analysis, we identified mutations associated with progression from CCUS to MDS by integrating mutation profiles, mutation frequencies and VAF trends.

2.7. Statistical analysis and visualization

The chi-square test was used to assess the statistical significance of the categorical variables, while continuous variables were compared using the Wilcoxon rank-sum test, as the data did not follow a normal distribution[31]. These statistical tests were performed using R (R Core Team, 2024) within the RStudio environment (Rstudio Team, 2024) and p-values < 0.05 were considered statistically significant in our study.

Mutation data were visualized and analyzed using the "Maftools" package in R[32]. Additionaly, we employed a custom "EasyOnco" script, which was developed using both Python and R, to generate MAF files from the result files of Piseq pipeline. The "EasyOnco" script was then utilized to create a customized oncoplot based on the "maftools" package.

3. Results

3.1. Patient characteristics

A total of 447 patients were categorized into the following groups: MDS, CCUS, ICUS, other hematologic malignancies (including acute myeloid leukemia (AML), chronic myeloid leukemia (CML), myelodysplastic/myeloproliferative neoplasms (MDS/MPN), and multiple myeloma (MM)), and other conditions (including solid tumors, metastasis, and nutritional deficiencies). As shown in **Figure 4**, 53.2% (n=238) of patients had MDS, and 11.0% (n=49) had CCUS, the two primary groups of focus in this study.

We compared baseline characteristics such as age, gender, bone marrow findings, and complete blood count (CBC) data as well as genetic characteristics between CCUS (n=49) and MDS (n=238) patients (**Table 2**). They showed significant differences in bone marrow blast (%) and hemoglobin level with p-value of <0.001 and 0.002, respectively. Dyspoiesis, including dysmegakaryopoiesis, dyserythropoiesis, and dysgranulopoiesis, as well as the blast percentage in bone marrow and peripheral blood, were higher in MDS, while hemoglobin (Hb) levels were lower.

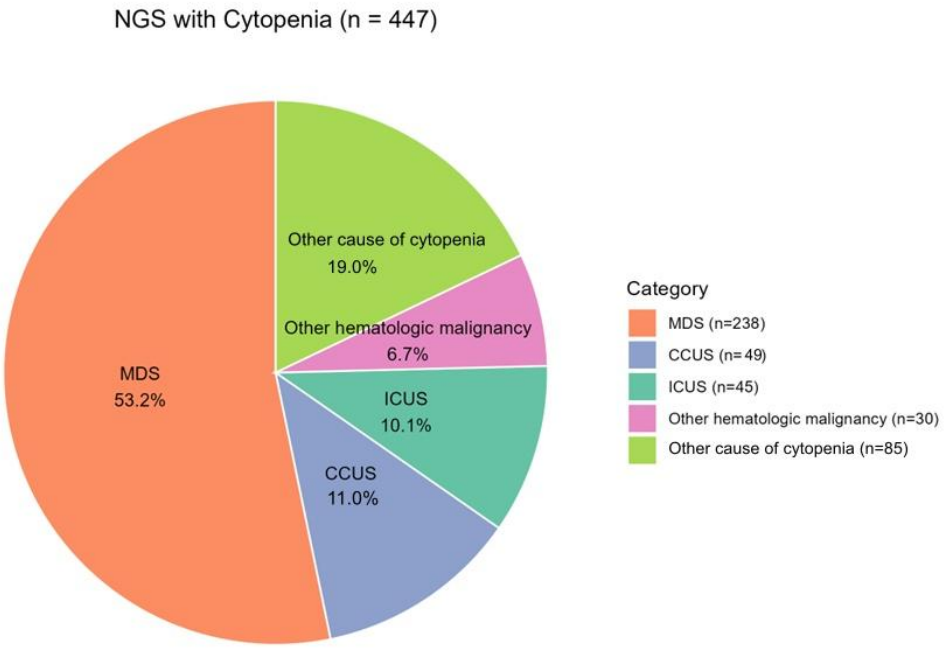

Figure 4. Distribution of patient groups in the study cohort.

Table 2. Clinical characteristics of MDS and CCUS patients enrolled in our study

	MDS (n=238)	CCUS (n=49)	p-val
Age(yr), mean(range)	66.0 (14-97)	68.1 (26-90)	0.354
Sex, n (%)			
Male	136 (57.4%)	27 (55.1%)	0.817
Female	101 (42.6%)	22 (44.9%)	
BM cellularity, n (%)			
Hypercellularity	87 (37.8%)	7 (14.3%)	
Normocellularity	97 (42.2%)	26 (53.1%)	0.246
Hypocellularity	46 (20.0%)	15 (30.6%)	
Dyspoiesis, mean(sd)			
Dysmegakaryopoiesis (%)	26.69 (27.10)	-	
Dyserythropoiesis (%)	11.26 (13.68)	-	
Dysgranulopoiesis (%)	7.31 (16.03)	-	
Ring Sideroblasts (%)	3.20 (11.54)	0.00 (0.00)	0.051
BM blast (%)	4.22 (4.69)	0.78 (0.70)	*<0.000
PB blast (%)	0.49 (1.63)	0.18 (1.32)	0.201
Complete Blood Count (CBC), mean(sd)			
White Blood Cell (WBC)	4.31 (8.96)	5.53 (10.82)	0.439
Absolute Neutrophil Count (ANC)	2.47 (8.01)	3.63 (10.22)	0.434
Haemoglobin (Hb)	8.54 (1.95)	9.63 (2.40)	*0.002
Platelet (PLT)	130.60 (127.87)	117.56 (98.99)	0.409

MDS, myelodysplastic syndrome; CCUS, clonal cytopenia of undetermined significance

*Bone marrow cellularity could not be evaluated in one CCUS case due to a dry-tap.

3.2. Genetic differences between MDS and CCUS

More than two thirds of patients with MDS had at least one tier 1/2 somatic mutation (178 out of 238 patients; 74.8%), whereas 95.9% of CCUS patients (47 out of 49) had at least one tier 1/2 somatic mutation (**Figure 5A**). The elevated proportion of mutated cases among CCUS patients is likely explained by the diagnostic criteria, which mandate the presence of at least one driver gene mutation or acquired chromosomal mosaicism. For the two CCUS patients without detected mutations, clonal chromosomal abnormalities or mutations in genes not included in the MRD30 panel were present that meet the CCUS diagnostic criteria. The number of somatic mutations per patient with at least one tier 1/2 mutation was significantly different between MDS and CCUS patients (**Figure 5B**). More specifically, MDS patients had a significantly higher average number of mutations compared to CCUS patients ($p = 0.0016$).

To further investigate the genetic differences between MDS and CCUS, we've looked at oncoplots based on tier 1/2 somatic mutations, including copy number variations (CNVs), identified in both groups. The oncoplot and its mutation summary for the MDS group (**Figure 6**, **Figure 7**) and the CCUS group (**Figure 8, Figure 9**) visually summarize the mutation frequencies and distributions across patient cohorts, highlighting the genes with the highest mutation rates in each group.

In MDS, *ASXL1* was the most frequently mutated gene, observed in 27.4% (n=46) of mutated patients, followed by *TP53* (26.2%, n=44), and *TET2* (25.6%, n=43). These three genes were the most common somatic mutations in MDS patients, suggesting their potential role in the pathogenesis of the disease.

In CCUS, *DNMT3A* was the most frequently mutated gene, seen in 44.9% (n=22) of mutated patients, followed by *TET2* (24.5%, n=12), and *TP53* (22.4%, n=11).

Table 3 summarizes the mutation frequencies of MRD30 panel-targeted genes in MDS and CCUS patients, along with the statistical significance of their differences. Considering the small sample size and the absence of mutations in certain genes within one of the groups, Fisher's exact test was used for statistical analysis. As you can see in **Figure 10**, the number of patients for *ASXL1, DNMT3A, EZH2, RUNX1* and *SF3B1* were significantly different between the two groups ($p < 0.05$). Notably, *ASXL1, EZH2, RUNX1* and *SF3B1* mutations were more prevalent in MDS, while *DNMT3A* mutations were more frequent in CCUS. Genes with significantly higher mutation frequencies in MDS may be associated with disease progression, while those predominantly mutated

in CCUS may serve as early indicators of clonal hematopoiesis.

The variant allele frequency (VAF%) of *TP53* was significantly different between the MDS and CCUS groups. The median VAF was 23.5% in MDS (n=238) and 0.55% in CCUS (n=49) for *TP53* (**Table 4**). Since the data did not meet the assumption of normality, we applied the Wilcoxon rank-sum test to get p-values. Although *TP53* did not exhibit significant differences in mutation frequency between the MDS and CCUS groups, the VAF analysis revealed statistically significant variations highlighting the importance of considering not only the presence of mutations but also the VAF values when comparing MDS and CCUS. These findings suggest that VAF values could provide additional insights into the molecular differences between these two conditions, which may not be captured by mutation frequency alone.

It is common for patients to have both germline and somatic mutations in *DDX41* gene, and several studies have examined the mutation sites of *DDX41*. Its mutation site is known to have differences between ethnicities[33]. As shown in the lollipop plot of *DDX41* in **Figure 11**, the most frequent mutated sites are p.V152G and p.Y259C. These mutations are likely hotspots due to their functional impact on the *DDX41* protein, which plays a crucial role in maintaining genomic stability. Previous studies have shown that the p.V152G mutation is particularly prevalent in the Korean population[34], but not in other ethnic groups, suggesting it may be a population-specific variant. Similarly, p.Y259C mutation is commonly found in both Korean and Japanese populations.

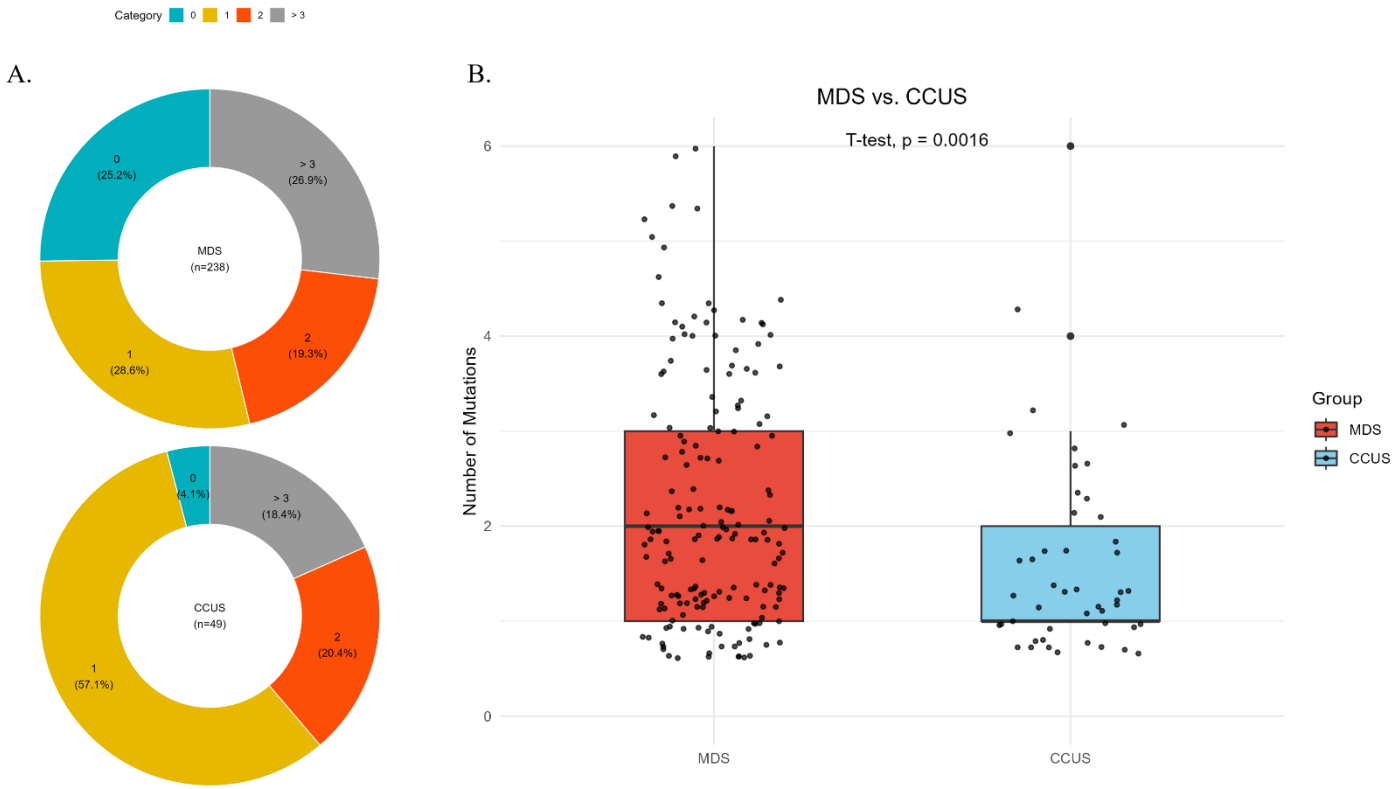

Figure 5. Distribution of tier 1/2 mutation in MDS and CCUS. (A) Donut charts shows the number of mutations in MDS and CCUS. The chart on the top represents MDS, while the one on the bottom represents CCUS. (B) The difference in mutation counts for individuals with at least on tier 1/2 mutation between MDS and CCUS (p-value = 0.0016)

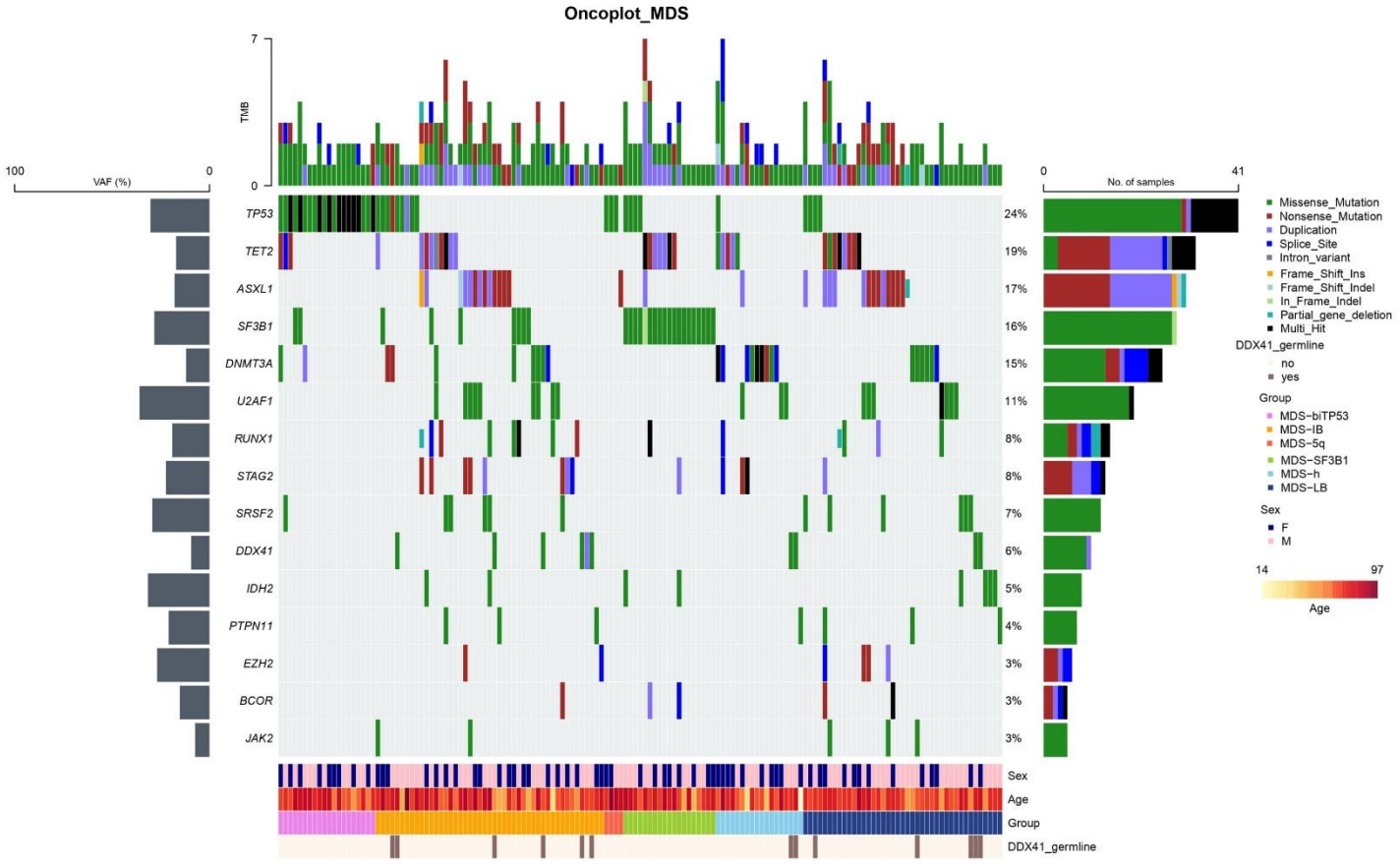

Figure 6. Oncoplot for MDS (n=238). The oncoplot above includes only cases harboring one or more mutations.

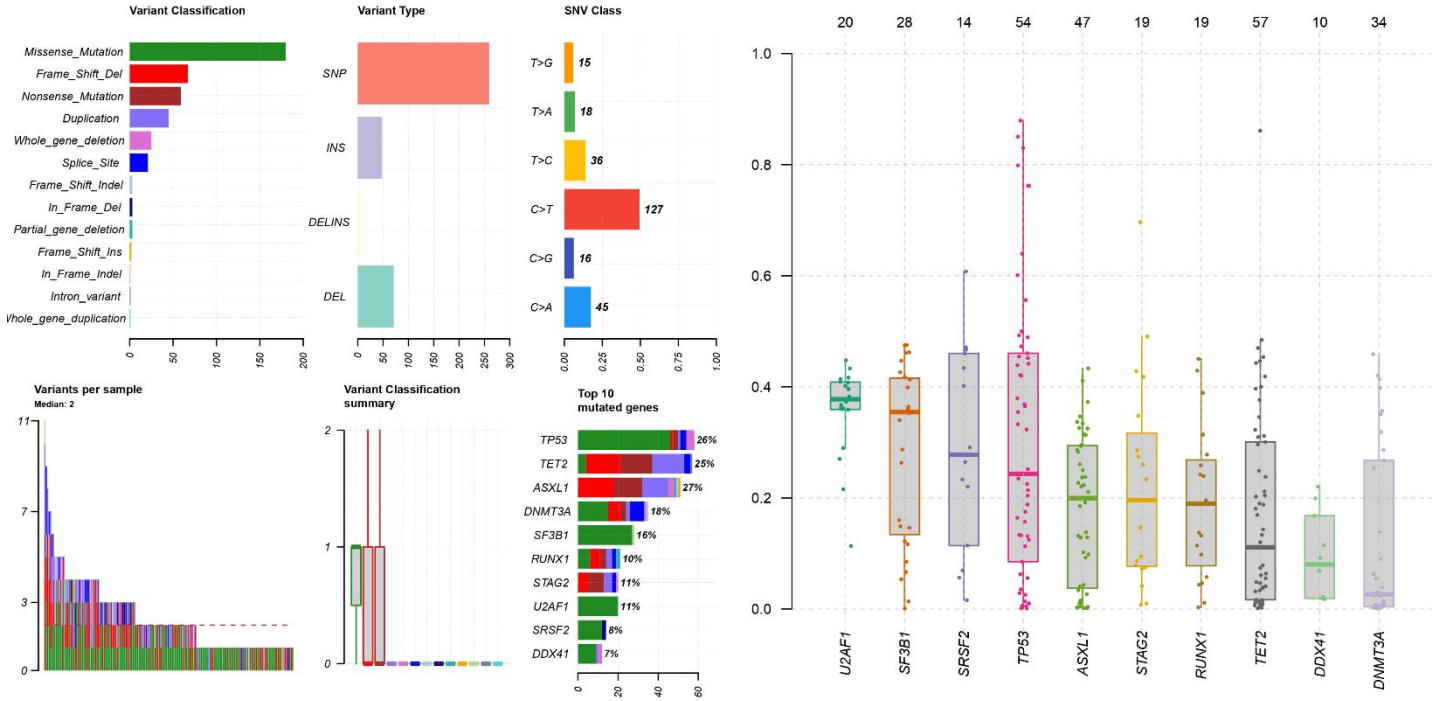

Figure 7. Summary of mutations in MDS. The left plot shows the summary of mutations in MDS, and the right plot shows the summary of VAF in each gene.

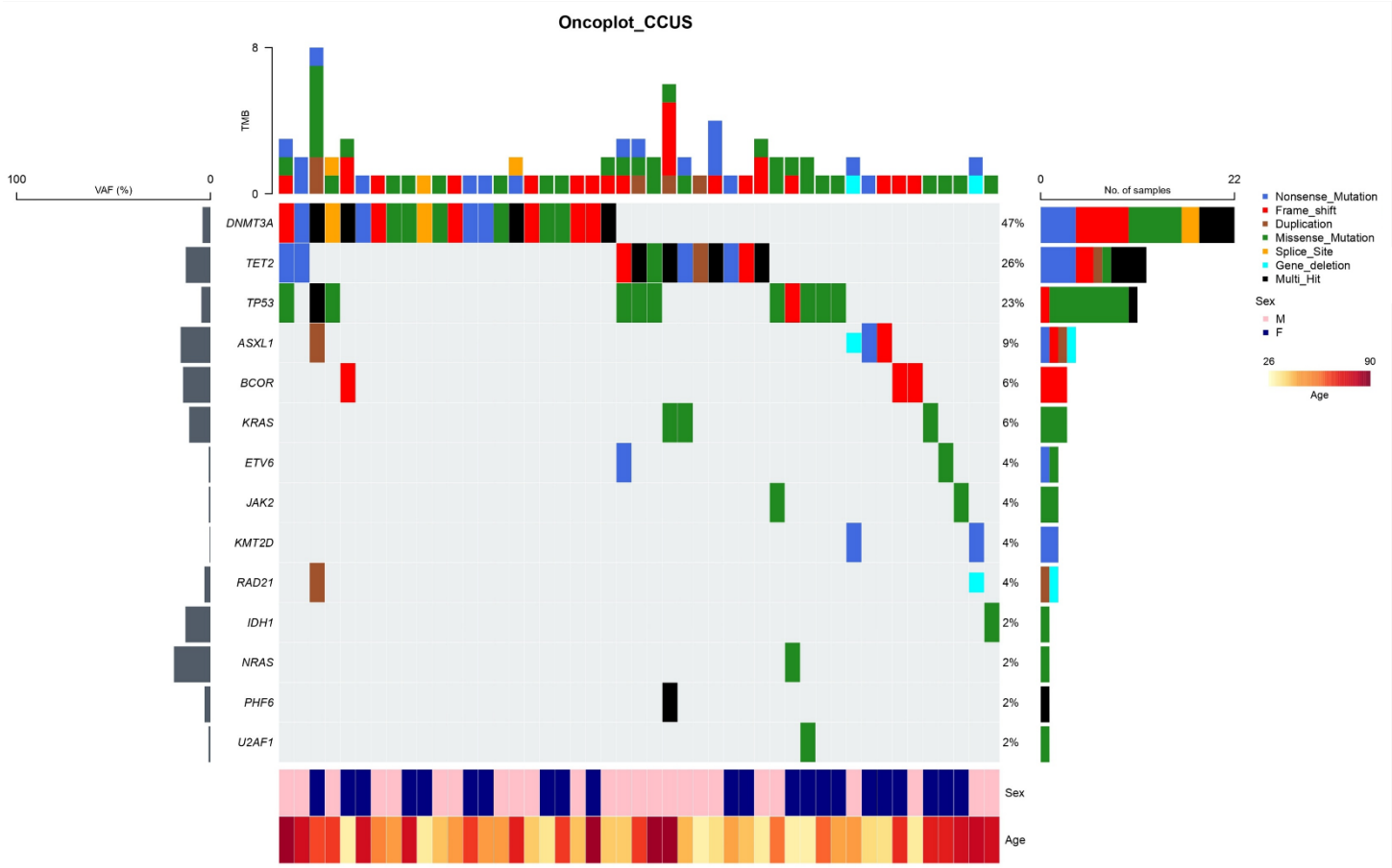

Figure 8. Oncoplot for CCUS (n=49). The oncoplot above includes only cases harboring one or more mutations.

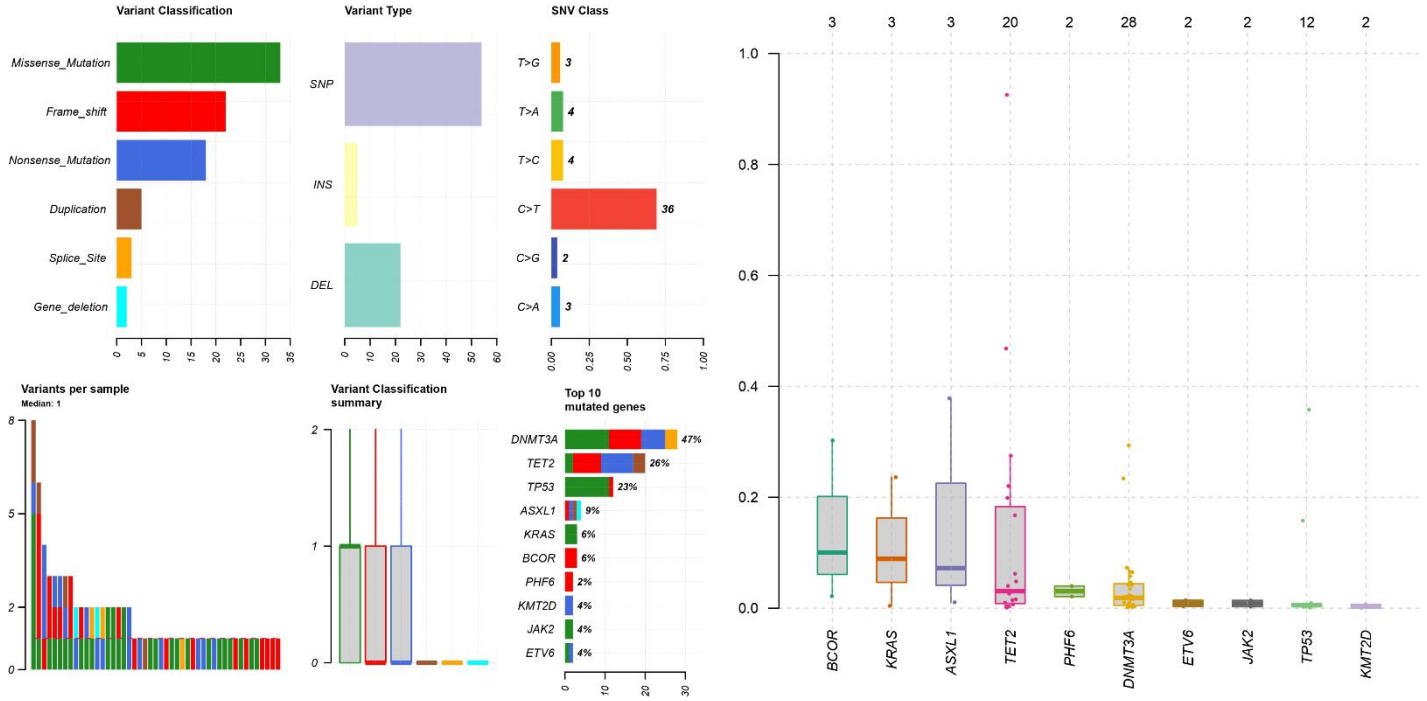

Figure 9. Summary of mutations in CCUS. The left plot shows the summary of mutations in CCUS, and the right plot shows the summary of VAF in each gene.

Table 3. Number of mutated patients in MDS and CCUS and the statistical significance of their differences

Gene	Number of patients (%)			Gene	Number of patients (%)		
	MDS (n=238)	CCUS (n=49)	p-value		MDS (n=238)	CCUS (n=49)	p-value
*ASXL1	47 (19.83%)	4 (8.16%)	0.029	NPM1	3 (1.27%)	0 (0.00%)	1.000
BCOR	8 (3.38%)	3 (6.12%)	0.440	NRAS	5 (2.11%)	1 (2.04%)	1.000
CEBPA	5 (2.11%)	0 (0.00%)	0.588	PHF6	5 (2.11%)	1 (2.04%)	1.000
DDX41	24 (10.13%)	1 (2.04%)	0.058	PTPN11	7 (2.95%)	0 (0.00%)	0.354
*DNMT3A	32 (13.08%)	21 (42.86%)	<0.001	RAD21	1 (0.42%)	2 (4.08%)	0.092
ETV6	1 (0.42%)	2 (4.08%)	0.092	*RUNX1	18 (7.59%)	0 (0.00%)	0.030
*EZH2	18 (7.59%)	0 (0.00%)	0.030	SETD2	2 (0.84%)	0 (0.00%)	1.000
IDH1	4 (1.69%)	1 (2.04%)	1.000	*SF3B1	28 (11.81%)	0 (0.00%)	0.004
IDH2	8 (3.38%)	0 (0.00%)	0.360	SRSF2	15 (6.33%)	0 (0.00%)	0.083
JAK2	5 (2.11%)	2 (4.08%)	0.176	*STAG2	19 (8.02%)	0 (0.00%)	0.030
KIT	2 (0.84%)	0 (0.00%)	1.000	TET2	43 (18.14%)	12 (24.49%)	0.567
KMT2D	6 (2.53%)	2 (4.08%)	0.648	TP53	44 (18.57%)	11 (22.45%)	0.849
KRAS	4 (1.69%)	3 (6.12%)	0.126	U2AF1	19 (8.02%)	1 (2.04%)	0.139
MPL	1 (0.42%)	0 (0.00%)	1.000	WT1	3 (1.27%)	0 (0.00%)	1.000

Variables marked with an asterisk(*) show statistical significance

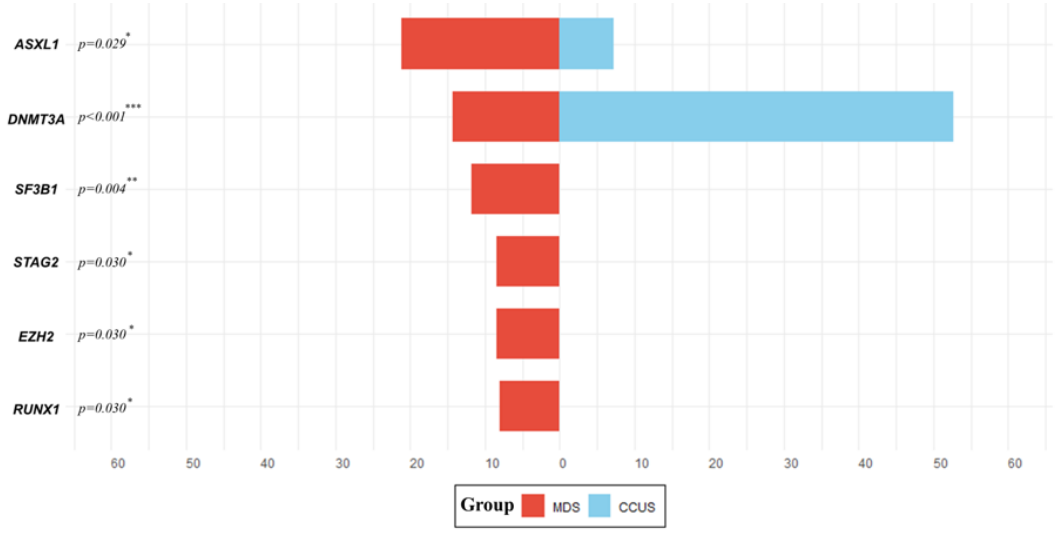

Figure 10. Percentage of patients with mutations in specific genes within MDS and CCUS cohorts. Only genes with statistically significant differences in mutation frequencies (p-value < 0.05) are shown.

Table 4. Comparison of gene-specific median VAF between MDS and CCUS

	Median VAF% (IQR)		
Gene	**MDS (n=238)**	**CCUS (n=49)**	**p-value**
****TP53***	23.49 (39.80)	0.53 (0.56)	<0.001
TET2	11.12 (28.87)	3.07 (18.33)	0.113
DNMT3A	2.99 (26.72)	1.87 (4.06)	0.203
KRAS	29.48 (45.45)	8.89 (23.19)	0.229
KMT2D	1.88 (3.23)	0.37 (0.70)	0.286
JAK2	1.78 (15.18)	1.40 (6.07)	0.549
PHF6	4.00 (27.60)	3.04 (3.99)	0.8
ASXL1	19.96 (26.11)	7.23 (36.79)	0.851
BCOR	7.18 (20.90)	10.03 (28.05)	1

Variables marked with an asterisk(*) show statistical significance

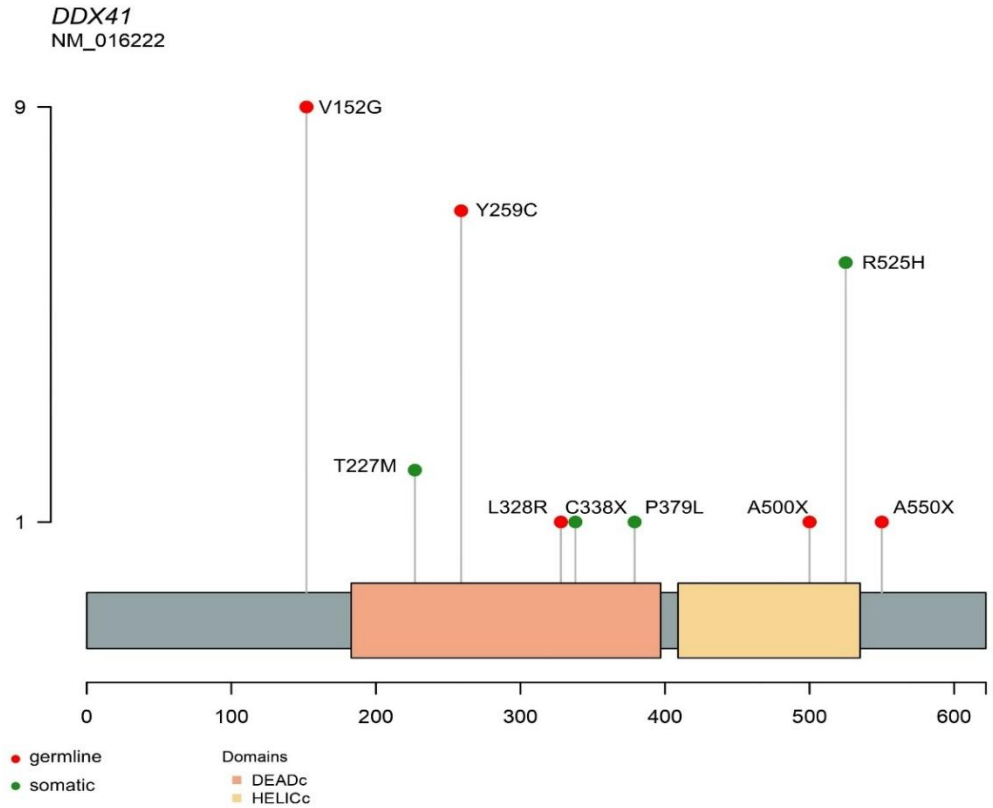

Figure 11. *DDX41* lollipop plot showing somatic and germline variants in MDS patients. The x-axis indicates amino acid positions, and the y-axis represents the number of mutations observed at each position. The most frequent mutation sites in *DDX41* in our cohort were p.V152G and p.Y259C, both of which are known to be common in the Korean population.

3.3. Investigating low VAF mutations without a cutoff

Unlike previous studies that analyzed mutations with VAF ≥ 2%, we removed the VAF cutoff to detect emerging clones at an earlier stage in patients who might have been overlooked. This approach could help identify potentially harmful variants before they reach higher VAF levels. Recent studies emphasize the importance of removing VAF threshold due to the variable fitness effects of each gene. Among our patients, 12 MDS and 20 CCUS patients had only low VAF mutations, with VAF under 2%.

We examined the mutated genes with low VAF, and *DNMT3A*, *TP53* and *TET2* were the most frequently mutated. *TET2* and *DNMT3A* are known to exhibit diverse fitness effects, making it essential to investigate even low VAF mutations[27].

For the six MDS subtypes according to the 2022 WHO classification, only three subtypes — MDS-LB, MDS-h, and MDS-IB—had samples classified as "low VAF only", with frequencies of 7.1%, 15.9%, and 6.2%, respectively (**Figure 12**). The other subtypes did not show a distinct "low VAF only" category. MDS-h exhibited the highest proportion of low VAF mutations. Notably, the 2022 WHO classification also describes that mutations in the MDS-h subtype are often detected at low VAFs, consistent with our findings [35]. Detecting these low-level mutations may aid in differentiating MDS-h from aplastic anemia. Given this characteristic, accurate diagnosis of MDS-h and similar subtypes requires sensitive detection methods capable of identifying low VAF mutations, underscoring the clinical importance of high-sensitivity sequencing approaches.

Some patients with low VAF mutations showed disease progression from CCUS to MDS and from MDS to AML. Patient shown in **figure 13** had a *TET2* c.294_297del mutation with a VAF of 2.6% at the time of CCUS diagnosis. Later, he was diagnosed with MDS, and the same *TET2* mutation was detected with a VAF of 2.78%. About six months later, the same *TET2* mutation was found with a VAF of 4.16%. Similarly, we observed two additional cases of disease progression from MDS to AML in patients who initially had low VAF mutations. One patient had a *TET2* c.4044+1G>C mutation with a VAF of 0.44% at the time of MDS diagnosis, and later, he was diagnosed to AML with the same *TET2* mutation at a VAF of 42.4% (**Figure 14**). The other patient had a *PTPN11* c.218C>T mutation with a VAF of 0.4% at the time of MDS diagnosis, and the same mutation was found with VAF of 34.6% when he was diagnosed with AML about a year later (**Figure 15**). These findings suggest that low VAF mutations in key genes may precede disease progression, highlighting the importance of monitoring such cases.

Among patients with only low VAF mutations (<2%), *DNMT3A*, *TP53*, *TET2*, and *ASXL1* were the most frequently mutated genes as shown on the oncoplot in **Figure 16**. These mutations are known for their roles in early clonal expansion and disease progression in hematologic malignancies. The presence of these mutations at low levels suggests that they may contribute to early clonal hematopoiesis or disease progression.

DNMT3A was the most frequently mutated gene, consistent with its role as an early clonal event in hematologic malignancies. *DNMT3A* mutations are commonly found in clonal hematopoiesis of indetermined significance (CHIP), CCUS, and early MDS, primarily representing early clonal expansion.

TP53 and *ASXL1* mutations, typically associated with poor prognosis and advanced disease, were also detected at low VAF in our study. The presence of subclonal *TP53* mutations even at low levels suggest that these alterations may exist before disease progression. Although we lack longitudinal follow-up data to confirm their clinical significance, these findings suggest that patients with low VAF *TP53* mutations may have an increased risk of progression, requiring close monitoring. Similarly, *ASXL1* mutations, which are known to be linked to poor prognosis and disease evolution in MDS, were also detected at low VAF, potentially serving as an early marker of clonal expansion and progression.

By removing the conventional VAF ≥ 2% cutoff, we were able to detect low VAF mutations that could have been missed. Our studies emphasize the clinical relevance of low VAF mutations in genes like *TP53*, *ASXL1*, *DNMT3A* and *TET2*, suggesting that they may serve as early indicators of disease progression. Future studies should focus on longitudinal monitoring to better understand the clinical significance of these mutations and their role in clonal evolution.

Among our cohort, three patients were confirmed to have progressed from CCUS to MDS based on serial bone marrow examinations. Notably, all three had a prior history of cytotoxic chemotherapy, raising the possibility that cytotoxic treatment may have contributed to clonal evolution.

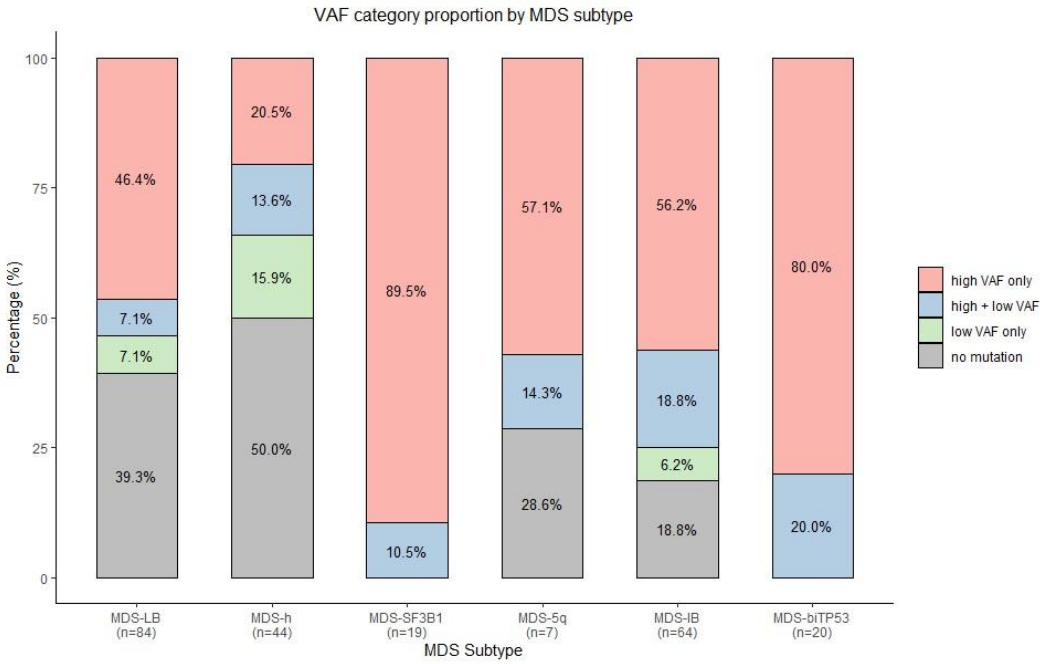

Figure 12. Distribution of mutation VAF categories by MDS subtypes. Samples are classified into four categories based on their mutation VAF: high VAF only (mutations with VAF ≥ 2%), high + low VAF (both high and low VAF mutations), low VAF only (mutations with VAF < 2%), and no mutation (no detected mutations). Sample sizes for each subtype are indicated on the x-axis labels. Among the subtypes, MDS-h showed the highest frequency (15.9%) of low VAF mutations.

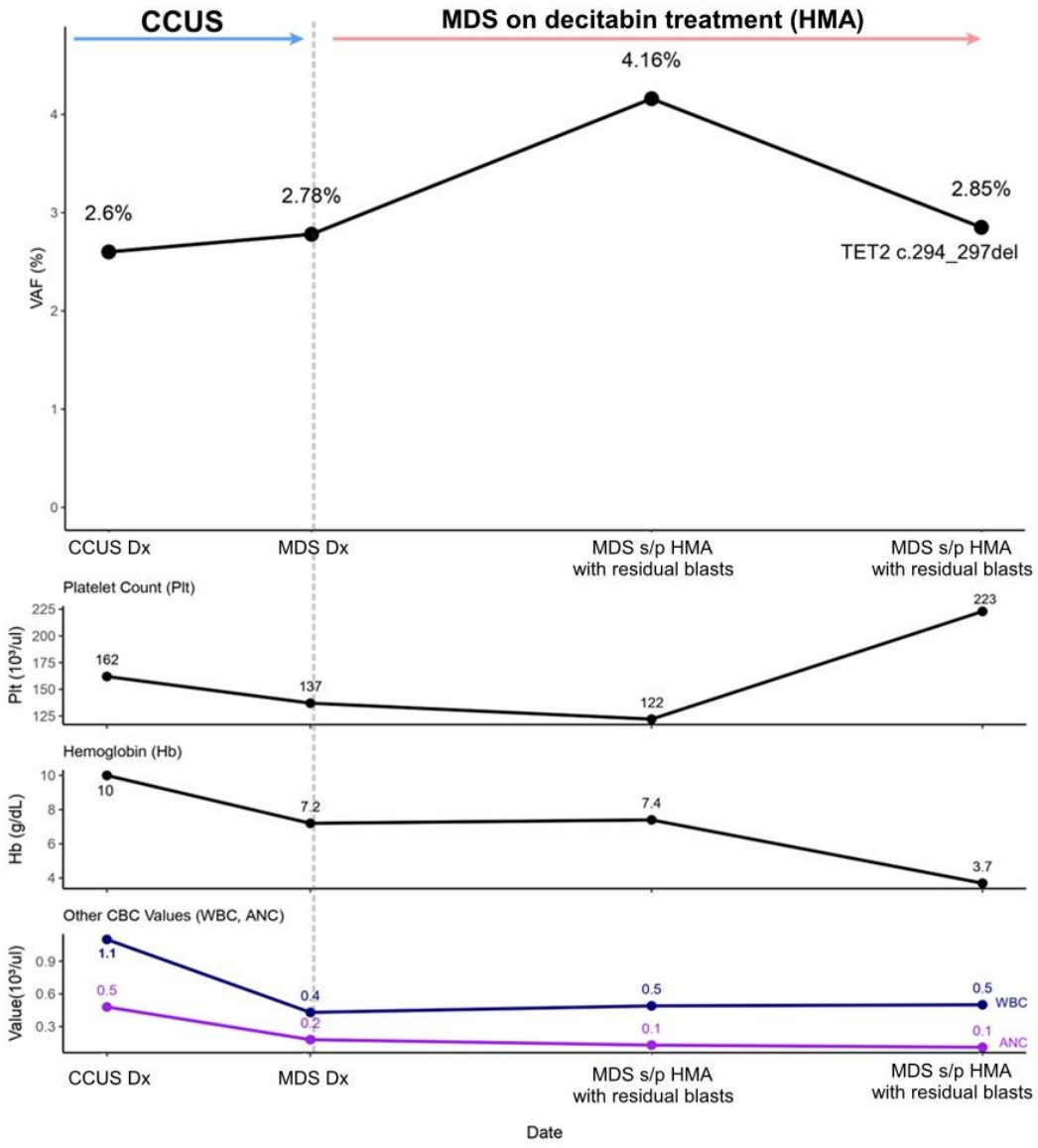

Figure 13. Clonal evolution in patients with low VAF mutations progressing from CCUS to MDS. This plot shows serial analysis of CBC trends, and mutational changes in patient progressing from CCUS to MDS.

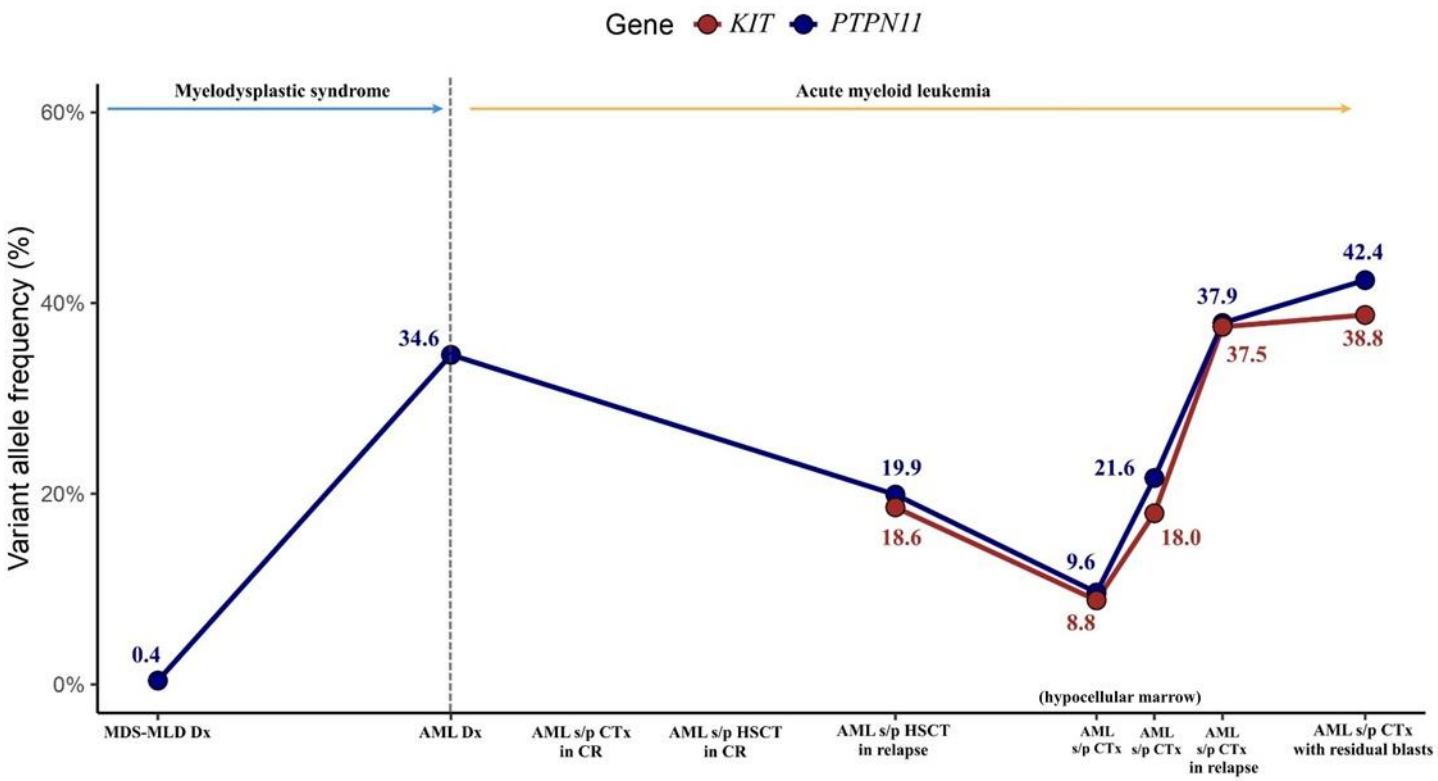

Figure 14. Clonal evolution in patients with low VAF mutations progressing from MDS to AML (case1). Abbreviations: MDS-MLD; myelodysplastic syndrome with multilineage dysplasia, Dx; diagnosis**,** CTx; chemotherapy, HSCT; hematopoietic stem cell transplant.

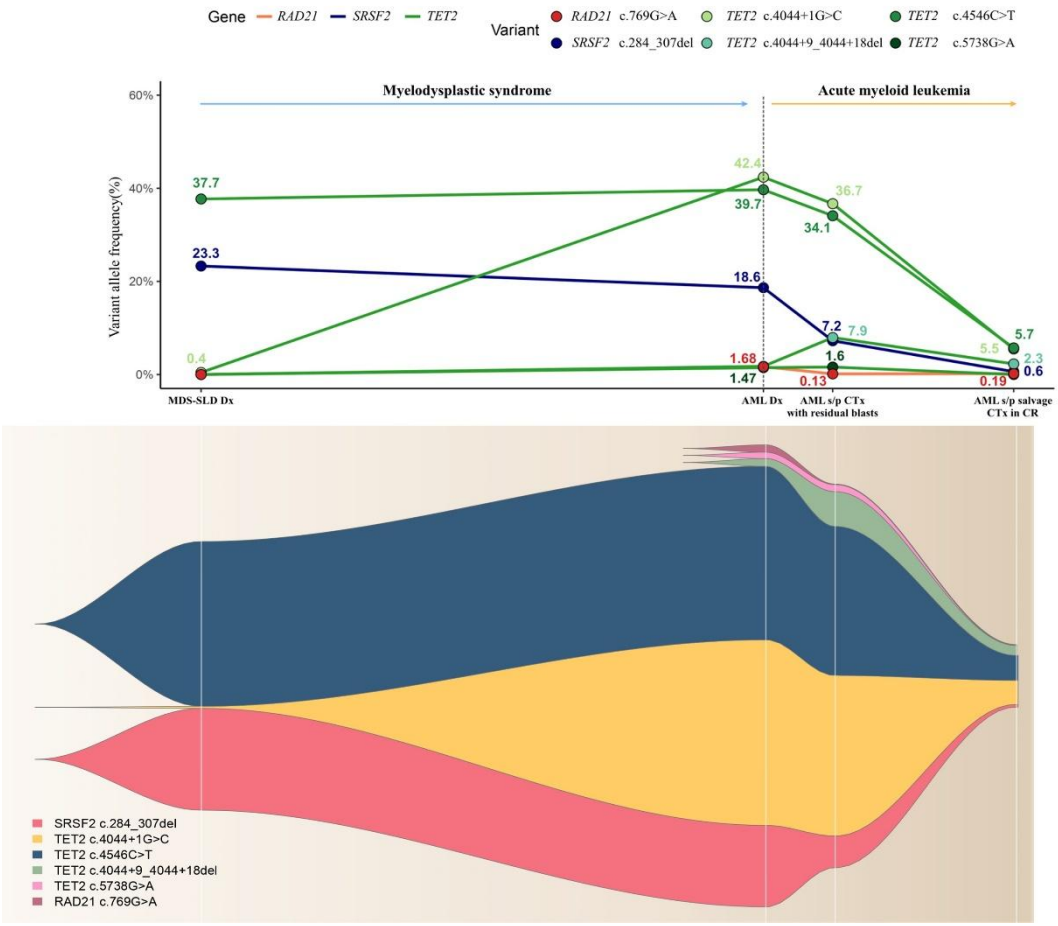

Figure 15. Clonal evolution in patients with low VAF mutations progressing from MDS to AML (case2).

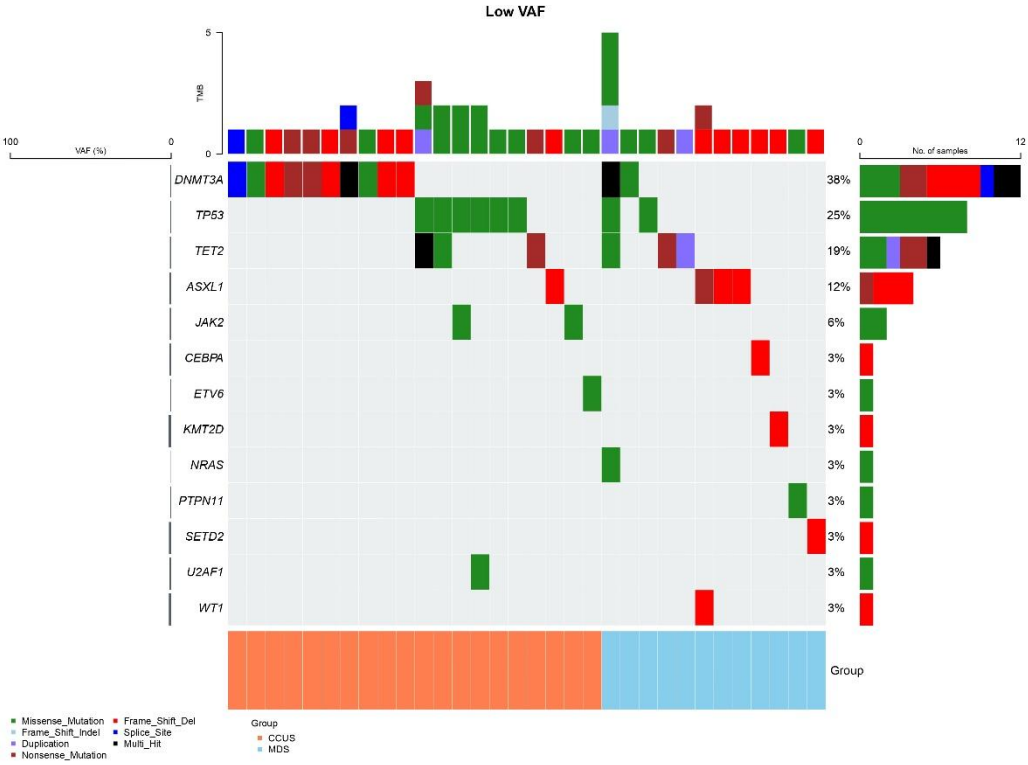

Figure 16. Oncoplot of patients with only low VAF mutations in MDS (n=12) and CCUS (n=20).

3.4. Functional categorization of mutated genes

To study the function of the mutated gene, we categorized genes into eight functional categories: DNA methylation (*DNMT3A, IDH1, IDH2, TET2*), tumor suppressors (*TP53, WT1, ETV6, PHF6*), activated signaling (*NRAS, FLT3, KIT, PTPN11, KRAS, JAK2, CALR, MPL*), chromatin modification (*ASXL1, EZH2, KMT2D*), transcription factors (*CEBPA, RUNX1, BCOR*), cohesion complexes (*STAG2, RAD21*), splicing factors (*SRSF2, U2AF1, SF3B1*), and others (*NPM1, SETD2, DDX41*)[36]. **Figure 17** shows the distribution and frequency of mutations categorized by gene function. The mutation frequency of DNA methylation-related genes was significantly higher in CCUS compared to MDS, while MDS showed higher mutation frequencies in splicing factors.

The increased mutation frequency of DNA methylation-related genes in CCUS is driven by the mutations in genes like *DNMT3A, TET2*, and *ASXL1*, which are well-established drivers of clonal hematopoiesis (CH)[37]. Particularly, *DNMT3A* and *TET2* are closely associated with hypermethylation, contributing to the higher mutation frequency observed in CCUS. This highlights the early-stage genetic alterations in CCUS that may predispose individuals to clonal hematopoiesis and increase the likelihood of progression to more severe hematologic conditions like MDS or AML.

In contrast, the higher mutation frequencies in splicing factors observed in MDS suggest that these mutations play a critical role in pathology of MDS. Splicing factor mutations disrupt RNA splicing and impair effective blood cell production, contributing to ineffective hematopoiesis, a hallmark of MDS[38]. The distinctive mutation patterns between CCUS and MDS suggest that while both disorders share certain genetic predispositions, the specific mutations present in each disease might underlie different disease mechanisms and progression pathways.

These findings emphasize the importance of genetic mutations in disease progression. In CCUS, mutations in DNA methylation-related genes could act as early indicators of clonal hematopoiesis, possibly identifying patients at higher risk for progression to MDS or AML. Meanwhile, the mutation patterns seen in MDS, particularly in splicing factors, underlie the complex genetic landscape that drives ineffective hematopoiesis and contributes to the disease's clinical presentation.

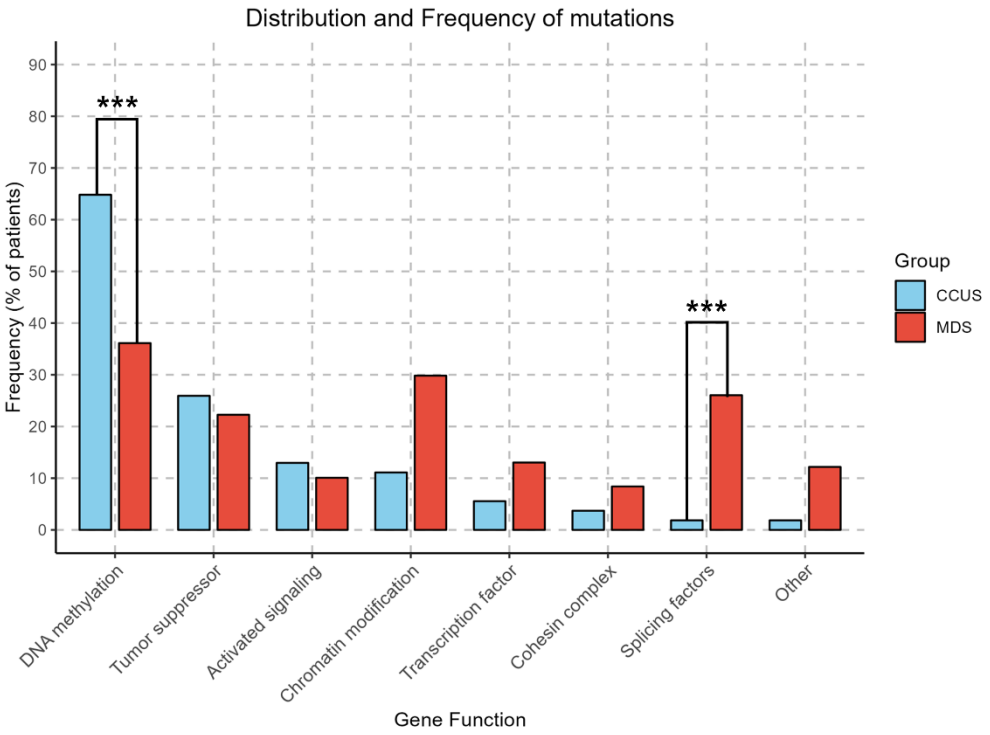

Figure 17. Distribution of mutations according to the functional categories in CCUS and MDS. This bar plot illustrates the distribution and frequency of gene mutations categorized by function in MDS and CCUS. DNA methylation-related genes (e.g., *DNMT3A, TET2, ASXL1*) show significantly higher mutation frequencies in CCUS ($p<0.001$), while splicing factor mutations (e.g., *SRSF2, U2AF1, SF3B1*) are more prevalent in MDS ($p<0.001$)[39], highlighting distinct genetic characteristics of each condition.

3.5. CBC Trends and Transfusion Dependency in CCUS

In patients diagnosed with CCUS, many were followed with serial complete blood counts (CBC) without immediate treatment. To investigate hematologic trends, we compared CBC changes between a “low VAF” group with maximum varint allele frequency < 2% and a “high VAF” group with maximum variant allele frequency ≥ 2%. As shown in **figure 18**, CBC values including white blood cell count (WBC), absolute neutrophil count (ANC), hemoglobin, and platelet count were plotted from the time of diagnosis through 3-6 months of follow-up.

Hemoglobin levels in the low VAF group remained consistently closer to the normal range compared to the high VAF group, and this difference was statistically significant ($p = 0.001$), as assessed using a linear mixed model to account for repeated measurements within individuals. Conversely, ANC and platelet counts tended to be closer to the normal range in the high VAF group. WBC counts showed no significant differences between the two groups.

An important clinical consideration in interpreting these trends is transfusion dependency. 40.9% of patients in the low VAF group and 65.4% in the high VAF group were TD. The apparent lack of a major difference in CBC trends between the groups may be influenced by this transfusion support. Moreover, some high VAF cases of CCUS may have been clinically treated as MDS based on cytopenias and physician judgement, which could further confound the comparison between groups.

To further explore the relationship between mutations and transfusion dependency, we compared mutational profiles between TD and transfusion independent (TID) patients (**Figure 19**). Although there was no statistical differences in mutations of TD and TID, median VAFs of *TET2* were slightly higher in the TD group than in the TID group.

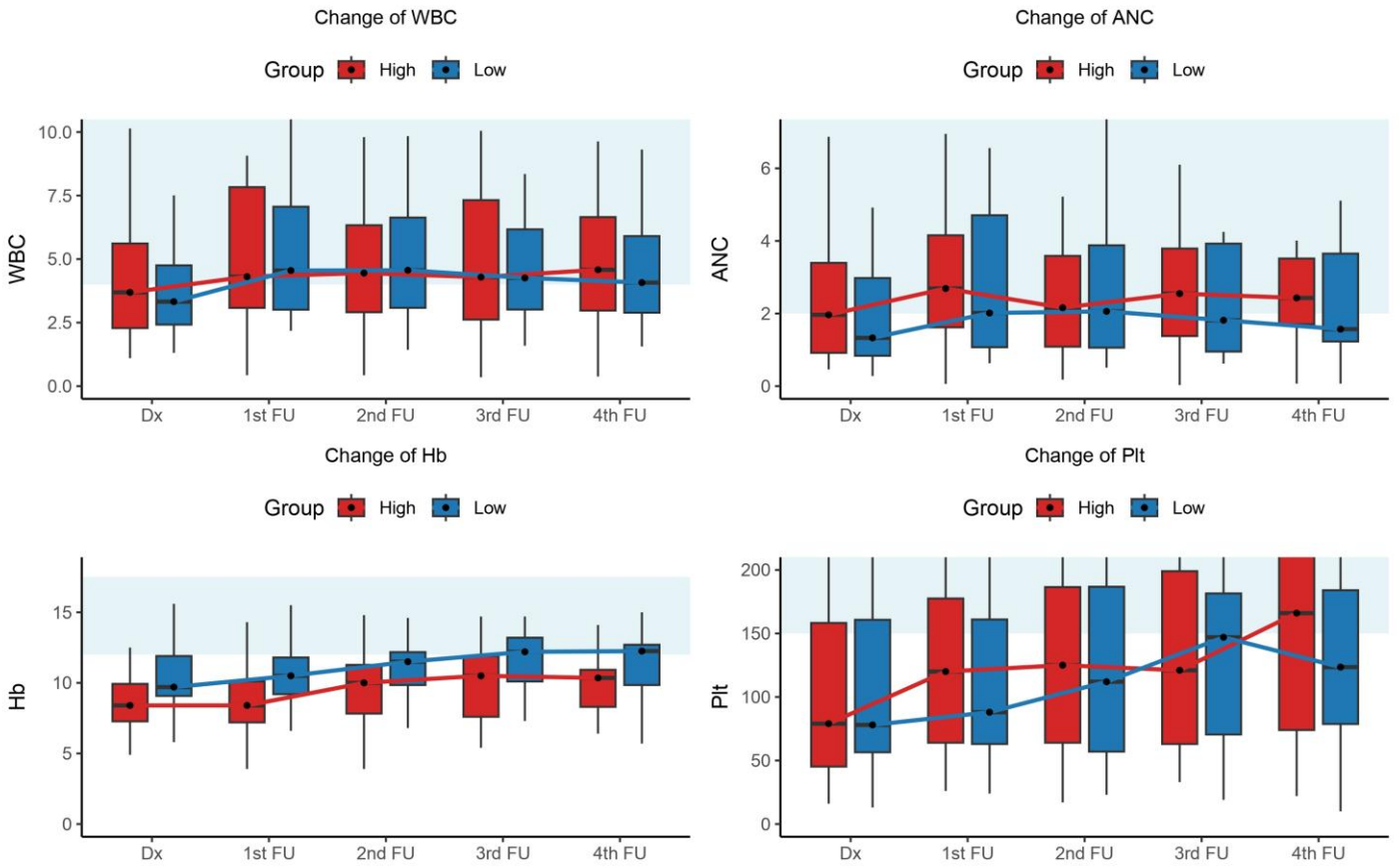

Figure 18. Longitudinal CBC trends in CCUS patients according to varint allele frequency. CBC values including WBC, ANC, hemoglobin, and platelets were compared between low (<2%) and high (≥2%) VAF groups at diagnosis and during 3-6 month follow ups. Shaded blue areas indicate normal reference ranges.

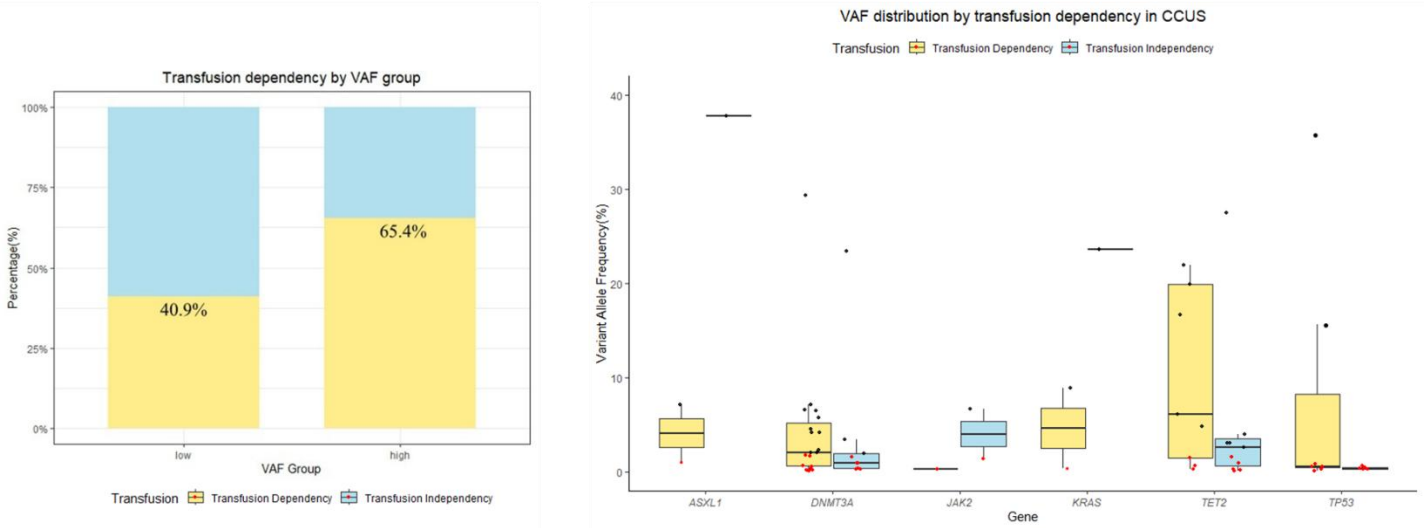

Figure 19. Mutation profiles in CCUS patients according to transfusion dependency status. Comparison of mutation patterns between transfusion dependent (TD) and transfusion independent (TID) CCUS patients. While no genes showed statistically significant differences, *TET2* mutations tended to have higher median VAFs in the TD group.

3.6. Meta-analysis

To enhance the reliability of our findings, we conducted a meta-analysis to integrate results from similar studies. **Figure 20** is the workflow of our meta-analysis based on the PRISMA guidelines. Following the literature review, five studies met our inclusion criteria[14,40,41,42,43]. Among the 30 genes in our MRD30 panel, four (*ASXL1, DNMT3A, SF3B1*, and *TET2*) were consistently included across all five comparative studies.

Figure 21A shows the meta-analysis result for *ASXL1*. In the pooled meta-analysis, *ASXL1* mutations did not show statistical significance, with the confidence interval (95% CI: -0.08 to 0.90) encompassing zero. However, in our study, the analysis using log (OR) revealed a statistically significant result, suggesting a potential association between *ASXL1* mutations and MDS.

Similarly, **Figure 21B** demonstrates results about *DNMT3A*. In the pooled meta-analysis, *DNMT3A* mutations were not statistically significant, with the confidence interval (95% CI: -1.23 to 0.41) including zero. However, in our study, a significant association was observed, with *DNMT3A* mutations being more frequently detected in CCUS compared to MDS. This finding suggests that *DNMT3A* mutations may play a more prominent role in the early stages of disease development, particularly in CCUS. The observed discrepancy between the pooled meta-analysis and our study highlights the importance of further research to better understand the role of *DNMT3A* mutations in the progression from CCUS to MDS. This difference may, in part, be explained by the application of VAF cutoffs in previous studies, which may exclude low-VAF mutations commonly seen in CCUS.

Figure 22 presents the meta-analysis results of other genes including *TET2*, *SF3B1*, *RUNX1*, *EZH2* and *STAG2*. For the genes not included in all five studies, we just analyzed them with available studies. While these genes showed a trend towards statistical significance in our study, the pooled analysis did not reach statistical significance, suggesting the need for further validation. The discrepancy between our study and the pooled meta-analysis data may be due to differences in cohort characteristics (e.g., some studies compared CCUS and LR-MDS, while others compared CCUS and MDS), VAF cutoffs, or sample sizes among studies.

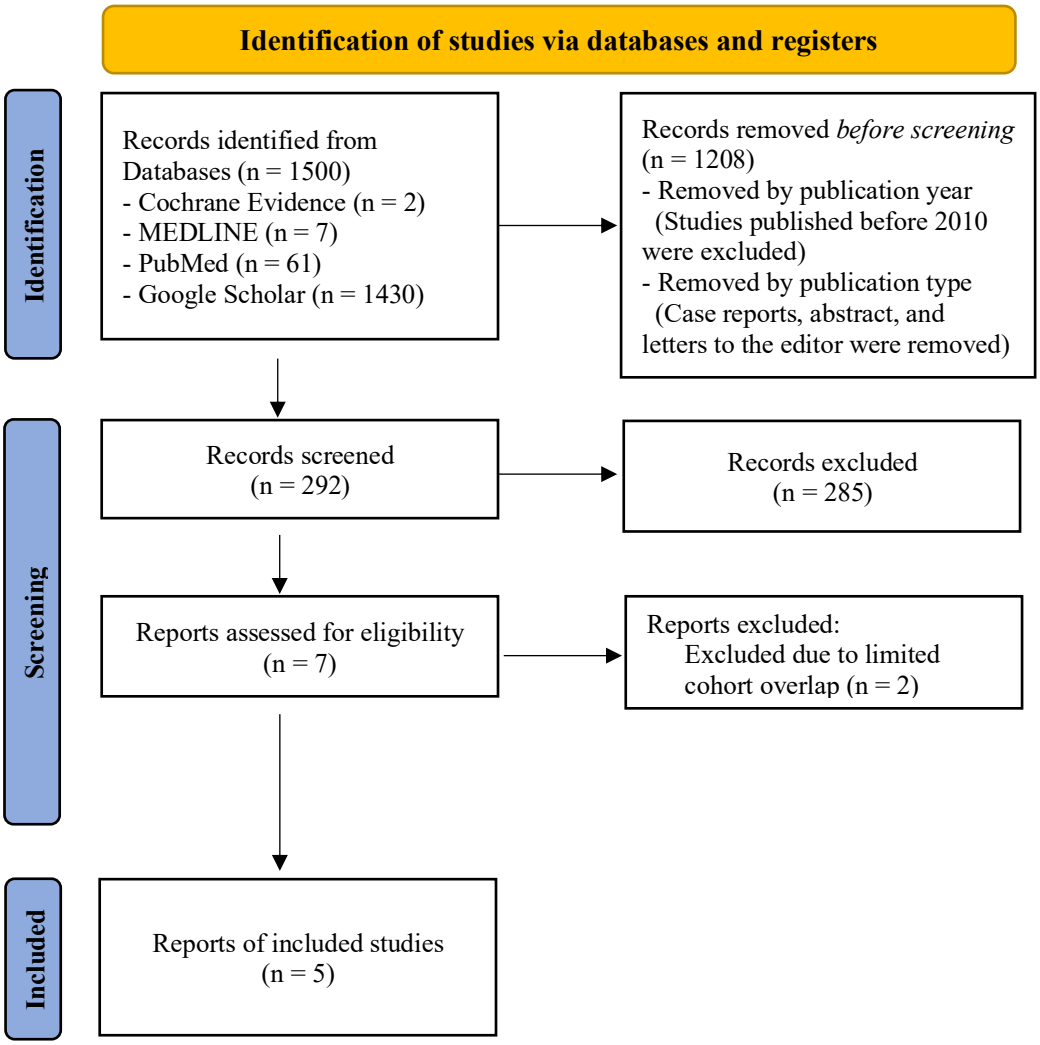

Figure 20. Meta-analysis workflow based on PRISMA 2020 flow diagram.

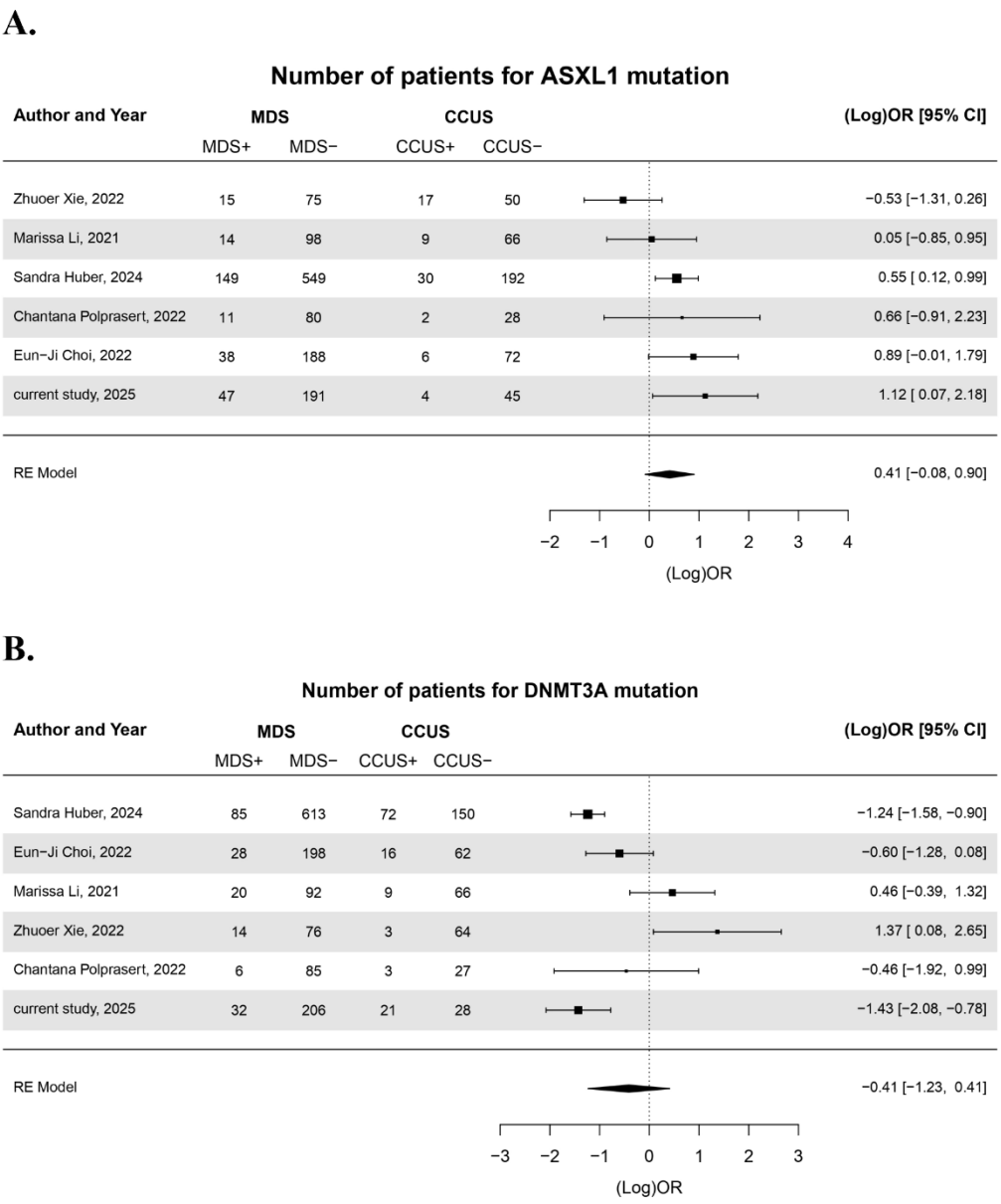

Figure 21. Meta-analysis results for *ASXL1* and *DNMT3A* mutations in the CCUS and MDS groups across five other studies.

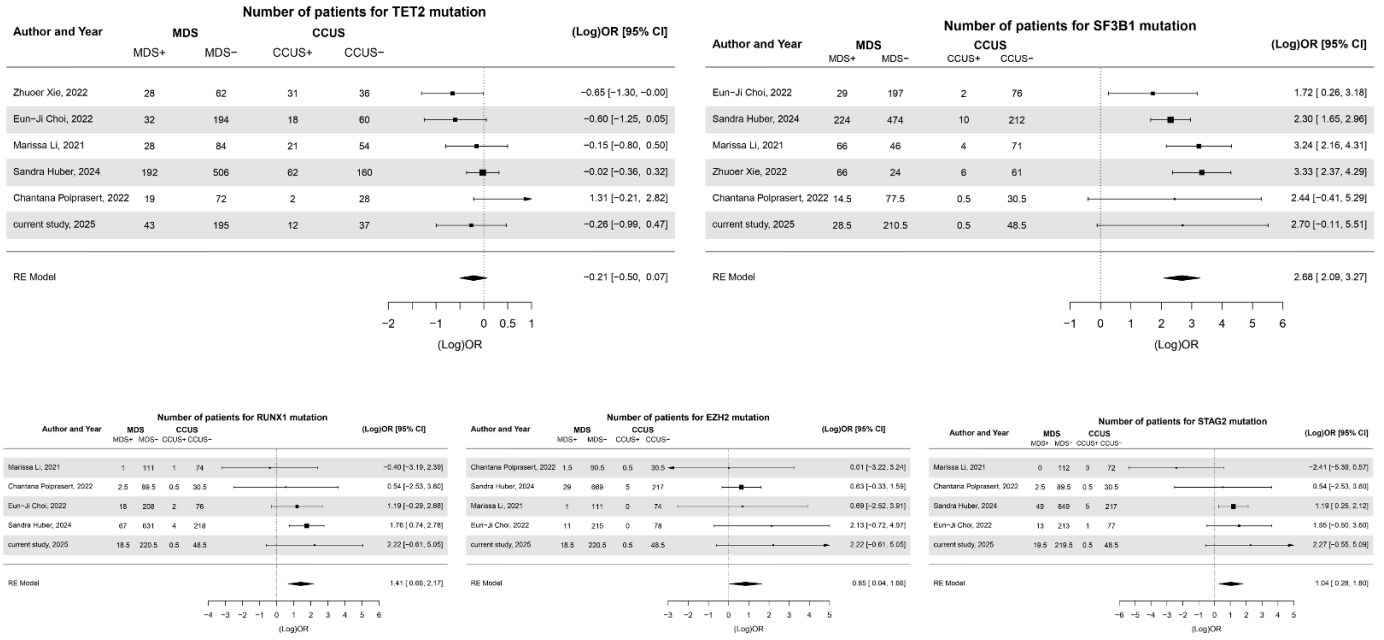

Figure 22. Meta-analysis of mutational impact in *TET2*, *SF3B1*, *RUNX1*, *EZH2*, and *STAG2*. While our study's results are not statistically significant, the trends were consistent with the overall meta-analysis results.

4. Discussion

This study aims to investigate the genetic differences between clonal hematopoiesis of undetermined significance (CCUS) and myelodysplastic syndrome (MDS), and to explore their implications for diagnosis and disease progression.

Our analysis demonstrated that MDS patients exhibited a higher somatic mutation burden compared to CCUS patients, with more than two-thirds of MDS patients harboring at least one tier 1/2 somatic mutation. The average number of mutations per patient was slightly higher in MDS (1.79) than in CCUS (1.57), and the average variant allele frequency (VAF) was also significantly higher in MDS (21.27%) compared to CCUS (7.22%). These findings reflect the more advanced clonal expansion and greater genetic complexity associated with MDS, highlighting the distinction between MDS and CCUS in terms of disease progression.

We also analyzed the mutated frequency of each gene to find the genetic differences between MDS and CCUS. We found that *ASXL1, SF3B1, RUNX1*, and *EZH2* were more commonly mutated in MDS, whereas *DNMT3A* mutations were predominant in CCUS. Functional categorization of the 30 genes in our MRD30 panel revealed distinct mutation patterns. DNA methylation-related mutations were significantly enriched in CCUS, whereas splicing factor mutations were more frequent in MDS. The predominance of DNA methylation-related mutations in CCUS suggests that epigenetic mechanisms may drive this precursor condition, potentially contributing to its development to MDS. In contrast, splicing factor mutations are closely linked with disease progression and unfavorable prognosis in MDS, highlighting their potential as biomarkers for monitoring MDS prognosis.

These findings have important clinical implications. The higher mutation burden and VAF observed in MDS patients suggest that they may have a higher risk of disease progression compared to CCUS. Understanding these genetic differences can help categorize patients based on their risk of progression from CCUS to MDS or other hematologic malignancies. The enrichment of DNA methylation-related mutations in CCUS points to pathogenic mechanism, potentially related to abnormal epigenetic regulation, which could inform future therapeutic strategies. One such strategy involves the use of hypomethylating agents (HMAs)[44], which are currently employed in MDS to

reverse abnormal DNA methylation and improve hematopoiesis. Given the potential role of DNA methylation in both CCUS and MDS, HMAs may be an appropriate therapeutic option for patients with methylation abnormalities.

In our study, we observed two cases of CCUS progressing to MDS in about a year. They each showed distinct mutational characteristics. The first case involved *PPM1D* mutation, which was not included in our MRD30 panel but was detected separately. This patient later progressed to therapy-related MDS (t-MDS), indicating the potential role of *PPM1D* mutations in therapy-related disease evolution. The second case showed an increasing VAF in *TET2* at the time of MDS diagnosis, showing the potential involvement of *TET2*-mutant clones in MDS development. These findings demonstrate the heterogeneous clonal dynamics of CCUS to MDS progression and suggest that specific mutations, especially those in *DNMT3A* and *TET2*, may contribute to disease evolution.

A limitation in this study lies in the relatively small number of patients who progressed to MDS. This can be attributed to the clinical approach used in the CCUS stage, where patients who exhibit clinical features similar to MDS are often treated early, even before a definitive diagnosis can be made. In clinical practice, when patients show signs resembling MDS, they are sometimes treated as if they have MDS, which can lead to an early initiation of treatment without performing additional bone marrow tests. As a result, while there is strong clinical suspicion of progression, the lack of follow-up bone marrow examinations prevents obtaining the definitive evidence needed to confirm MDS progression. Additionally, many of the patients in this study were elderly, and some declined further treatment or follow-up, leading to a lack of additional bone marrow evaluations and follow-up data.

Moreover, this study was conducted at a single center, which may limit the generalizability of the findings. The follow-up duration was relatively short, and only a targeted panel of 30 genes was analyzed, which may not fully capture the broader genomic landscape relevant to disease progression.

For meta-analysis, further studies with larger sample sizes and standardized methods are necessary to validate the role of interested genes in early clonal evolution. Moreover, longitudinal studies tracking VAF changes over time could provide deeper understanding of how these mutations affect disease progression.

5. Conclusion

In conclusion, our study explores the genetic differences between CCUS and MDS, emphasizing the key mutations and molecular patterns that can distinguish these two conditions. By identifying specific genetic markers, such as the predominance of *DNMT3A* mutations in CCUS (found in 40% of patients) and splicing factor mutations in MDS (present in 26.1% of MDS patients), we provide valuable insights into their pathophysiology and the risk of disease progression. Our findings highlight the importance of early detection and treatment of CCUS, which may help delay or prevent the transition from CCUS to MDS in certain patients.

However, there were some challenges in confirming disease progression, especially due to the lack of follow-up bone marrow evaluations in elderly patients or those who declined further treatment. This emphasizes the need for innovative approaches in both diagnosis and patient care, as well as more comprehensive follow-up protocols to better track disease progression over time.

Also, we discovered that applying the commonly used 2% variant allele frequency (VAF) cutoff may result in undetected mutations for certain genes. For example, in our study, 64.5% of mutations in *DNMT3A* and 46.8% of mutations in *TET2* would have been missed using the 2% cutoff. This suggests that adjusting the VAF threshold for individual genes could enable more sensitive distinctions between CCUS and MDS, improving the accuracy of diagnosis and potentially aiding in more precise risk stratification.

While our study contributed to understanding the genetics of clonal hematopoiesis and its role in hematologic malignancies, further research is necessary. Larger cohorts and trackable extended follow-up periods will be crucial in refining our understanding of the genetic and clinical factors that drive disease progression. Such research will also be necessary for the development of improved diagnostic tools and prognostic models, which will lead to better patient outcomes and more effective management of these complex conditions.

REFERENCES

1. Kang Y, Shin S, Choi Y, Lee S, Choi J, Kook H, et al. 772P Comparative analysis of myelodysplastic syndrome (MDS) and clonal cytopenia of undetermined significance (CCUS) using a genetic approach. Annals of Oncology 2024;35:S1689.
2. Zink F, Stacey SN, Norddahl GL, Frigge ML, Magnusson OT, Jonsdottir I, et al. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. Blood, The Journal of the American Society of Hematology 2017;130:742-52.
3. Karner K, George TI, Patel JL. Current aspects of clonal hematopoiesis: implications for clinical diagnosis. Annals of Laboratory Medicine 2019;39:509-14.
4. Gaulin C, Kelemen K, Arana Yi C. Molecular pathways in clonal hematopoiesis: from the acquisition of somatic mutations to transformation into hematologic neoplasm. Life 2022;12:1135.
5. Bowman RL, Busque L, Levine RL. Clonal hematopoiesis and evolution to hematopoietic malignancies. Cell stem cell 2018;22:157-70.
6. Bewersdorf JP, Xie Z, Bejar R, Borate U, Boultwood J, Brunner AM, et al. Current landscape of translational and clinical research in myelodysplastic syndromes/neoplasms (MDS): Proceedings from the 1st International Workshop on MDS (iwMDS) Of the International Consortium for MDS (icMDS). Blood Reviews 2023;60:101072.
7. Jaiswal S, Ebert BL. Clonal hematopoiesis in human aging and disease. Science 2019;366:eaan4673.
8. Bewersdorf JP, Ardasheva A, Podoltsev NA, Singh A, Biancon G, Halene S, et al. From clonal hematopoiesis to myeloid leukemia and what happens in between:

Will improved understanding lead to new therapeutic and preventive opportunities? Blood reviews 2019;37:100587.

9. Valent P. ICUS, IDUS, CHIP and CCUS: diagnostic criteria, separation from MDS and clinical implications. Pathobiology 2019;86:30-8.
10. DeZern AE, Malcovati L, Ebert BL. CHIP, CCUS, and other acronyms: definition, implications, and impact on Practice. American Society of Clinical Oncology Educational book. American Society of Clinical Oncology. Annual Meeting; 2019. p.400-10.
11. Valent P, Horny H-P, Bennett JM, Fonatsch C, Germing U, Greenberg P, et al. Definitions and standards in the diagnosis and treatment of the myelodysplastic syndromes: Consensus statements and report from a working conference. Leukemia research 2007;31:727-36.
12. Xie Z, Komrokji R, Al Ali N, Regelson A, Geyer S, Patel A, et al. Risk prediction for clonal cytopenia: multicenter real-world evidence. Blood 2024;144:2033-44.
13. Khanna V, Lu R, Kumar J, Molina A, Stehr H, Spiteri E, et al. The clinical, molecular, and prognostic features of the 2022 WHO and ICC classification systems for myelodysplastic neoplasms. Leukemia Research 2024;136:107433.
14. Li M, Binder M, Lasho T, Ferrer A, Gangat N, Al-Kali A, et al. Clinical, molecular, and prognostic comparisons between CCUS and lower-risk MDS: a study of 187 molecularly annotated patients. Blood Advances 2021;5:2272-8.
15. Ogawa S. Genetics of MDS. Blood, The Journal of the American Society of Hematology 2019;133:1049-59.
16. Gorczyca W. Atlas of differential diagnosis in neoplastic hematopathology: CRC Press; 2021.
17. Falini B, Martelli MP. Comparison of the International Consensus and 5th WHO edition classifications of adult myelodysplastic syndromes and acute myeloid leukemia. American journal of hematology 2023;98:481-92.

18. Khoury JD, Solary E, Abla O, Akkari Y, Alaggio R, Apperley JF, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. Leukemia 2022;36:1703-19.
19. Xie Z, Nanaa A, Saliba AN, He R, Viswanatha D, Nguyen P, et al. Treatment outcome of clonal cytopenias of undetermined significance: a single-institution retrospective study. Blood cancer journal 2021;11:43.
20. Li H. Toward better understanding of artifacts in variant calling from high-coverage samples. Bioinformatics 2014;30:2843-51.
21. Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, Loeb LA. Detection of ultra-rare mutations by next-generation sequencing. Proceedings of the National Academy of Sciences 2012;109:14508-13.
22. Lee KS, Seo J, Lee C-K, Shin S, Choi Z, Min S, et al. Analytical and clinical validation of cell-free circulating tumor DNA assay for the estimation of tumor mutational burden. Clinical chemistry 2022;68:1519-28.
23. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. The Journal of molecular diagnostics 2017;19:4-23.
24. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine 2015;17:405-23.
25. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in bioinformatics 2013;14:178-92.

26. Kim JJ, Jang JE, Lee HA, Park MR, Kook HW, Lee S-T, et al. Development of a next-generation sequencing-based gene panel test to detect measurable residual disease in acute myeloid leukemia. Annals of Laboratory Medicine 2023;43:328-36.
27. Robertson NA, Latorre-Crespo E, Terradas-Terradas M, Lemos-Portela J, Purcell AC, Livesey BJ, et al. Longitudinal dynamics of clonal hematopoiesis identifies gene-specific fitness effects. Nature medicine 2022;28:1439-46.
28. Bannon SA, Routbort MJ, Montalban-Bravo G, Mehta RS, Jelloul FZ, Takahashi K, et al. Next-generation sequencing of DDX41 in myeloid neoplasms leads to increased detection of germline alterations. Frontiers in oncology 2021;10:582213.
29. Winstone L, Jung Y, Wu Y. DDX41: exploring the roles of a versatile helicase. Biochemical Society Transactions 2024;52:395-405.
30. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. BMJ Ment Health 2019;22:153-60.
31. Mishra P, Pandey CM, Singh U, Keshri A, Sabaretnam M. Selection of appropriate statistical methods for data analysis. Annals of cardiac anaesthesia 2019;22:297-301.
32. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome research 2018;28:1747-56.
33. Li P, Brown S, Williams M, White T, Xie W, Cui W, et al. The genetic landscape of germline DDX41 variants predisposing to myeloid neoplasms. Blood, The Journal of the American Society of Hematology 2022;140:716-55.
34. Choi E-J, Cho Y-U, Hur E-H, Jang S, Kim N, Park H-S, et al. Unique ethnic features of DDX41 mutations in patients with idiopathic cytopenia of undetermined significance, myelodysplastic syndrome, or acute myeloid leukemia. Haematologica 2021;107:510.

35. Bono E, McLornan D, Travaglino E, Gandhi S, Gallì A, Khan AA, et al. Clinical, histopathological and molecular characterization of hypoplastic myelodysplastic syndrome. Leukemia 2019;33:2495-505.
36. Menezes J, Acquadro F, Wiseman M, Gómez-López G, Salgado R, Talavera-Casanas J, et al. Exome sequencing reveals novel and recurrent mutations with clinical impact in blastic plasmacytoid dendritic cell neoplasm. Leukemia 2014;28:823-9.
37. Pich O, Reyes-Salazar I, Gonzalez-Perez A, Lopez-Bigas N. Discovering the drivers of clonal hematopoiesis. Nature communications 2022;13:4267.
38. Boultwood J, Dolatshad H, Varanasi SS, Yip BH, Pellagatti A. The role of splicing factor mutations in the pathogenesis of the myelodysplastic syndromes. Advances in biological regulation 2014;54:153-61.
39. Brunner AM, Leitch HA, van de Loosdrecht AA, Bonadies N. Management of patients with lower-risk myelodysplastic syndromes. Blood cancer journal 2022;12:166.
40. Huber S, Baer C, Hutter S, Wossidlo N, Hoermann G, Pohlkamp C, et al. Genomic landscape of CCUS compared to MDS and its implications on risk prediction. Leukemia 2024;38:1634-7.
41. Choi EJ, Cho YU, Hur EH, Park HS, Choi Y, Lee JH, et al. Clinical implications and genetic features of clonal cytopenia of undetermined significance compared to lower-risk myelodysplastic syndrome. British Journal of Haematology 2022;198:703-12.
42. Xie Z, Campestri G, Lasho T, Finke C, Li M, Binder M, et al. Clonal compositions involving epigenetic regulator and splicing mutations in CHIP, CCUS, MDS, and CMML. Leukemia research 2022;116:106818.
43. Polprasert C, Niparuck P, Rattanathammethee T, Chuncharunee S, Kobbuaklee S, Songserm K, et al. Excellent Prognosis of Low-Risk Myelodysplastic Syndromes

(MDS) Without Detectable Myeloid-Related Mutations. Clinical Lymphoma Myeloma and Leukemia 2022;22:e293-e9.

44. Santini V. How I treat MDS after hypomethylating agent failure. Blood, The Journal of the American Society of Hematology 2019;133:521-9.

ABSTRACT IN KOREAN

골수이형성증후군(MDS)과 원인 불명 클론성 혈구감소증(CCUS) 에서의 유전적 특성 비교

미확정의 클론성 조혈(CCUS)은 원인불명의 지속적인 혈구 감소증과 클론성 조혈(CH)이 동반되는 질환으로, 골수이형성 증후군(MDS)과 감별 진단이 필요하다. CCUS는 현재 명확한 치료 기준이 확립되지 않았으며, 일부 환자는 장기적으로 MDS로 진행할 가능성이 있어 조기 진단과 예후 예측이 중요하다. 그러나 CCUS와 MDS는 임상적, 혈액학적 유사성이 있어 감별이 어려우며, 특히 유전적 특성에 대한 이해가 부족한 실정이다. 이에 본 연구에서는 CCUS와 MDS의 유전적 특징을 분석하여 두 질환의 병태생리적 차이를 규명하고, 임상적 함의를 도출하고자 하였다.

본 연구에서는 지속적인 혈구 감소증을 보이는 환자들을 대상으로 차세대 염기서열 분석(NGS)을 시행하여 유전자 변이 양상을 비교하였다. 이를 통해 CCUS와 MDS에서 관찰되는 주요 유전자 변이 패턴을 분석하고, 각 질환의 병리 기전에 미치는 영향을 평가하였다. 특히, DNA 메틸화 관련 유전자 및 스플라이싱 인자 변이의 분포 차이를 확인함으로써 CCUS의 발병 기전과 MDS로의 진행 가능성을 이해하는 데 초점을 맞추었다. 또한 각 질환에서 관찰된 유전적 변이의 빈도와 변이 대립유전자 빈도(VAF)를 비교함으로써 클론 확장의 차이를 분석하고, 이러한 차이가 질병 진행 및 위험도 평가에 미치는 영향을 탐색하였다.

본 연구는 CCUS와 MDS의 유전적 차이를 밝힘으로써 두 질환의 감별 진단을 위한 새로운 근거를 제공하며, 향후 예후 예측 모델 개발 및 치료 전략 수립에 기여할 수 있을 것으로 기대된다. 이러한 연구를 통해 MDS로의 진행 전에 CCUS의 진단 및 치료에 대한 새로운 방향을 제시할 수 있으며, 환자 개개인에게 최적화된 관리 방법을 제공하는 데 기여하고자 한다.

핵심되는 말: 골수이형성증후군, 미확정 클론성 조혈, 차세대 염기서열 분석